

**PATENT APPLICATION**

**METHODS AND APPARATUSES FOR ANALYZING  
POLYNUCLEOTIDE SEQUENCES**

Inventors:

Stephen Quake

Wayne Volkmuth

Marc Unger

Assignee:

Mycometrix Corporation  
213 E. Grand Avenue  
South San Francisco, CA 94080

Entity:

Small

sub A1

## METHODS AND APPARATUSES FOR ANALYZING POLYNUCLEOTIDE SEQUENCES

5

### CROSS-REFERENCES TO RELATED APPLICATIONS

This nonprovisional patent application claims the benefit of the previously  
filed patent applications: U.S. provisional patent application no. 60/163,742, filed November  
4, 1999; and U.S. patent application no. 09/605,520, filed June 27, 2000, which in turn claims  
10 the benefit of U.S. provisional patent application no. 60/141,503 filed June 28, 1999, U.S.  
provisional patent application no. 60/147,199 filed August 3, 1999, and U.S. provisional  
patent application no. 60/186,856 filed March 3, 2000. The text of these previously filed  
patent applications is hereby incorporated by reference.

15

### STATEMENT AS TO RIGHTS TO INVENTIONS MADE UNDER FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

Work described herein has been supported, in part, by NIH grants HG-01642-  
02. The U. S. Government may therefore have certain rights in the invention.

20

### TECHNICAL FIELD

The present invention relates to methods for high speed, high throughput  
analysis of polynucleotide sequences and apparatuses for carrying out such methods.

25

### BACKGROUND OF THE INVENTION

Traditional DNA sequencing techniques share three essential steps in their  
approaches to sequence determination. First, a multiplicity of DNA fragments are generated  
from a DNA species which it is intended to sequence. These fragments are incomplete copies  
30 of the DNA species to be sequenced. The aim is to produce a ladder of DNA fragments, each  
a single base longer than the previous one. For example, with the Sanger method (Sanger et  
al., Proc. Natl. Acad. Sci. USA 74:5463, 1977), the target DNA is used as a template for a  
DNA polymerase to produce a number of incomplete clones. These fragments, which differ

in respective length by a single base, are then separated on an apparatus which is capable of resolving single-base differences in size. The third and final step is the determination of the nature of the base at the end of each fragment. When ordered by the size of the fragments which they terminate, these bases represent the sequence of the original DNA species.

5 Automated systems for DNA sequence analysis have been developed, such as discussed in Toneguzzo et al., 6 Biotechniques 460, 1988; Kanbara et al., 6 Biotechnology 816, 1988; and Smith et al., 13 Nuc. Acid. Res. 13: 2399, 1985; U.S. Pat. No. 4,707,237 (1987). However, all these methods still require separation of DNA products by a gel permeation procedure and then detection of their locations relative to one another along the  
10 axis of permeation or movement through the gel. These apparatuses used in these methods are not truly automatic sequencers. They are merely automatic gel readers, which require the standard sequencing reactions to be carried out before samples are loaded onto the gel.

The disadvantages of the above methods are numerous. The most serious problems are caused by the requirement for the DNA fragments to be size-separated on a  
15 polyacrylamide gel. This process is time-consuming, uses large quantities of expensive chemicals, and severely limits the number of bases which can be sequenced in any single experiment, due to the limited resolution of the gel. Sanger dideoxy sequencing has a read length of approximately 500 bp, a throughput that is limited by gel electrophoresis (appropriately 0.2%).

20 Other methods for analyzing polynucleotide sequences have been developed more recently. In some of these methods broadly termed sequencing by synthesis, template sequences are determined by priming the template followed by a series of single base primer extension reactions (e.g., as described in WO 93/21340, WO 96/27025, and WO 98/44152). While the basic scheme in these methods no longer require separation of polynucleotides on  
25 the gel, they encounter various other problems such as consumption of large amounts of expensive reagents, difficulty in removing reagents after each step, misincorporation due to long exchange times, the need to remove labels from the incorporated nucleotide, and difficulty to detect further incorporation if the label is not removed. Many of these difficulties stem directly from limitations of the macroscopic fluidics employed. However,  
30 small-volume fluidics have not been available. As a result, these methods have not replaced the traditional gel-based sequencing schemes in practice. The skilled artisans are to a large extent still relying on the gel-based sequencing methods.

Thus, there is a need in the art for methods and apparatuses for high speed and high throughput analysis of longer polynucleotide sequences which can be automated

utilizing the available scanning and detection technology. The present invention fulfills this and other needs.

## SUMMARY OF THE INVENTION

5

In one aspect of the present invention, methods for analyzing the sequence of a target polynucleotide are provided. The methods include the steps of (a) providing a primed target polynucleotide linked to a microfabricated synthesis channel; (b) flowing a first nucleotide through the synthesis channel under conditions whereby the first nucleotide  
10 attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide; (c) determining presence or absence of a signal, the presence of a signal indicating that the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; (d) removing or reducing the signal, if present; and (e) repeating steps (b)-(d) with a further  
15 nucleotide that is the same or different from the first nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.

In some methods, step (a) comprises providing a plurality of different primed target polynucleotides linked to different synthesis channels; step (b) comprises flowing the first nucleotide through each of the synthesis channels; and step (c) comprises determining  
20 presence or absence of a signal in each of the channels, the presence of a signal in a synthesis channel indicating the first nucleotide was incorporated into the primer in the synthesis channel, and hence the identity of the complementary base that served as a template in the target polynucleotide in the synthesis channel. In some methods, a plurality of different primed target polynucleotides are linked to each synthesis channels.

25 Some methods include the further steps of flushing the synthesis channel to remove unincorporated nucleotides. In some methods, steps (b)-(d) are performed at least four times with four different types of nucleotides. In some methods, steps (b)-(d) are performed until the identity of each base in the target polynucleotide has been identified.

In some methods, the nucleotides are labeled. The label can be a fluorescent  
30 dye, and the signal can be detected optically. The label can also be a radiolabel, and the signal can be detected with a radioactivity detector. In some methods, incorporation of nucleotides is detected by measuring pyrophosphate release.

In some methods, the synthesis channel is formed by bonding a microfluidic chip to a flat substrate. In some of these methods, the target polynucleotides are immobilized

to the interior surface of the substrate in the synthesis channel. In some of these methods, the interior surface is coated with a polyelectrolyte multilayer (PEM). In some of these methods, the microfluidic chip is fabricated with an elastomeric materia such as RTV silicone.

In another aspect of the present invention, methods for analyzing a target polynucleotide entails (a) pretreating the surface of a substrate to create surface chemistry that facilitates polynucleotide attachment and sequence analysis; (b) providing a primed target polynucleotide attached to the surface; (c) providing a labeled first nucleotides to the attached target polynucleotide under conditions whereby the labeled first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide; (d) determining presence or absence of a signal from the primer, the presence of a signal indicating that the labeled first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and (e) repeating steps (c)-(d) with a labeled further nucleotide that is the same or different from the first labeled nucleotide, whereby the labeled further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.

In some of these methods, the substrate is glass and the surface is coated with a polyelectrolyte multilayer (PEM). In some methods, the PEM is terminated with a polyanion. In some methods, the polyanion is terminated with carboxylic acid groups. In some methods, the target polynucleotide is biotinylated, and the PEM-coated surface is further coated with biotin and then streptavidin.

In still another aspect of the present invention, methods of analyzing a target polynucleotide are provided which include the steps of (a) providing a primed target polynucleotide; (b) providing a first type of nucleotide of which a fraction is labeled under conditions whereby the first nucleotide attaches to the primer, if a complementary nucleotide is present to serve as template in the target polynucleotide; (c) determining presence or absence of a signal from the primer, the presence of a signal indicating the first nucleotide was incorporated into the primer, and hence the identity of the complementary base that served as a template in the target polynucleotide; and (d) repeating steps (b)-(c) with a further type of nucleotide of which a fraction is labeled the same and which is the same or different from the first type of nucleotide, whereby the further nucleotide attaches to the primer or a nucleotide previously incorporated into the primer.

In some of these methods, the label used is a fluorescent label. In some of these methods, the removing or reducing step is performed by photobleaching. In some of

these methods, the fraction of labeled nucleotides are less than 10%, less than 1%, less than 0.1%, or less than 0.01%.

In another aspect of the present invention, apparatuses for analyzing the sequence of a polynucleotide are provided. The apparatuses have (a) a flow cell with at least one microfabricated synthesis channel; and (b) an inlet port and an outlet port which are in fluid communication with the flow cell and which flowing fluids such as deoxynucleoside triphosphates and nucleotide polymerase into and through the flow cell. Some of the apparatuses additionally have (c) a light source to direct light at a surface of the synthesis channel; and (d) a detector to detect a signal from the surface.

In some of the apparatuses, the synthesis channel is formed by bonding a microfluidic chip to a flat substrate. In some apparatuses, the microfluidic chip also contain microfabricated valves and microfabricated pumps in an integrated system with the synthesis channel. In some of these apparatuses, a plurality of reservoirs for storing reaction reagents are also present, and the microfabricated valve and pump are connected to the reservoirs. In some apparatuses, the detector is a photon counting camera. In some of the apparatuses, the microfluidic chip is fabricated with an elastomeric material such as RTV silicone. The substrate of some of the apparatuses is a glass cover slip. The cross section of the synthesis channel in some of the apparatuses has a linear dimension of less than  $100\mu\text{m} \times 100\mu\text{m}$ , less than  $10\mu\text{m} \times 100\mu\text{m}$ , less than  $1\mu\text{m} \times 10\mu\text{m}$ , or less than  $0.1\mu\text{m} \times 1\mu\text{m}$ .

## BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is an illustration of a first elastomeric layer formed on top of a micromachined mold.

Fig. 2 is an illustration of a second elastomeric layer formed on top of a micromachined mold.

Fig. 3 is an illustration of the elastomeric layer of Fig. 2 removed from the micromachined mold and positioned over the top of the elastomeric layer of Fig. 1

Fig. 4 is an illustration corresponding to Fig. 3, but showing the second elastomeric layer positioned on top of the first elastomeric layer.

Fig. 5 is an illustration corresponding to Fig. 4, but showing the first and second elastomeric layers bonded together.

Fig. 6 is an illustration corresponding to Fig. 5, but showing the first micromachine mold removed and a planar substrate positioned in its place.

Fig. 7A is an illustration corresponding to Fig. 6, but showing the elastomeric structure sealed onto the planar substrate.

Figs. 7B is a front sectional view corresponding to Fig. 7A, showing an open flow channel.

5 Fig. 7C corresponds to Fig. 7A, but shows a first flow channel closed by pressurization in second flow channel.

Fig. 8 is an illustration of a first elastomeric layer deposited on a planar substrate.

10 Fig. 9 is an illustration showing a first sacrificial layer deposited on top of the first elastomeric layer of Fig. 8.

Fig. 10 is an illustration showing the system of Fig. 9, but with a portion of the first sacrificial layer removed, leaving only a first line of sacrificial layer.

15 Fig. 11 is an illustration showing a second elastomeric layer applied on top of the first elastomeric layer over the first line of sacrificial layer of Fig. 10, thereby encasing the sacrificial layer between the first and second elastomeric layers.

Fig. 12 corresponds to Fig. 11, but shows the integrated monolithic structure produced after the first and second elastomer layers have been bonded together.

Fig. 13 is an illustration showing a second sacrificial layer deposited on top of the integral elastomeric structure of Fig. 12.

20 Fig. 14 is an illustration showing the system of Fig. 13, but with a portion of the second sacrificial layer removed, leaving only a second line of sacrificial layer.

Fig. 15 is an illustration showing a third elastomer layer applied on top of the second elastomeric layer and over the second line of sacrificial layer of Fig. 14, thereby encapsulating the second line of sacrificial layer between the elastomeric structure of Fig. 12 and the third elastomeric layer.

Fig. 16 corresponds to Fig. 15, but shows the third elastomeric layer cured so as to be bonded to the monolithic structure composed of the previously bonded first and second elastomer layers.

30 Fig. 17 corresponds to Fig. 16, but shows the first and second lines of sacrificial layer removed so as to provide two perpendicular overlapping, but not intersecting, flow channels passing through the integrated elastomeric structure.

Fig. 18 is an illustration showing the system of Fig. 17, but with the planar substrate thereunder removed.

Fig. 19 illustrates valve opening vs. applied pressure for various flow channels.

Fig. 20 illustrates time response of a  $100\mu\text{m}\times 100\mu\text{m}\times 10\mu\text{m}$  RTV microvalve.

Fig. 21 is a schematic illustration of a multiplexed system adapted to permit flow through various channels.

Fig. 22A is a plan view of a flow layer of an addressable reaction chamber structure.

Fig. 22B is a bottom plan view of a control channel layer of an addressable reaction chamber structure.

Fig. 22C is an exploded perspective view of the addressable reaction chamber structure formed by bonding the control channel layer of Fig. 22B to the top of the flow layer of Fig. 22A.

Fig. 22D is a sectional elevation view corresponding to Fig. 22C, taken along line 28D-28D in Fig. 22C.

Fig. 23 is a schematic of a system adapted to selectively direct fluid flow into any of an array of reaction wells.

Fig. 24 is a schematic of a system adapted for selectable lateral flow between parallel flow channels.

Fig. 25 is a schematic of an integrated system for analyzing polynucleotide sequences.

Fig. 26 is a schematic of a further integrated system for analyzing polynucleotide sequences.

Fig. 27 is a schematic diagram of a sequencing apparatus.

## 25 DETAILED DESCRIPTION

### I. Overview

The present invention provides methods and apparatuses for analyzing polynucleotide sequences.

In some methods, the sequencing apparatuses comprise a microfabricated flow channel to which polynucleotide templates are attached. Optionally, the apparatuses comprise a plurality of microfabricated channels, and diverse polynucleotide templates can be attached to each channel. The apparatuses can also have a plurality of reservoirs for storing



various reaction reagents, and pumps and valves for controlling flow of the reagents. The flow cell can also have a window to allow optical interrogation.

In these methods, single stranded polynucleotide templates with primers are immobilized to the surface of the microfabricated channel or to the surface of reaction chambers that are disposed along a microfabricated flow channel, e.g., with streptavidin-biotin links. After immobilization of the templates, a polymerase and one of the four nucleotide triphosphates are flowed into the flow cell, incubated with the template, and flowed out. If no signal is detected, the process is repeated with a different type of nucleotide.

These methods are advantageous over the other sequencing by synthesis methods discussed previously. First, use of microfabricated sequencing apparatuses reduces reagent consumption. It also increases reagent exchange rate and the speed of sequence analysis. In addition, the microfabricated apparatuses provides parallelization: many synthesis channels can be built on the same substrate. This allows analysis of a plurality of diverse polynucleotide sequences simultaneously. Further, due to the reduction of time and dead volume for exchanging reagents between different steps during the analysis, mismatch incorporation is greatly reduced. Moreover, the read length is also improved because there is less time for the polymerase to incorporate a wrong nucleotide and it is less likely that the polymerase falls off the template. All these advantages result in high speed and high throughput sequence analysis regimes.

In some methods of the present invention, the surface of a substrate (e.g., a glass cover slip) is pretreated to create optimal surface chemistry that facilitates polynucleotide template attachment and subsequent sequence analysis. In some of these methods, the substrate surface is coated with a polyelectrolyte multilayer (PEM). Following the PEM coating, biotin can be applied to the PEM, and followed by application of streptavidin. The substrate surface can then be used to attach biotinylated-templates. The PEM-coated substrate provides substantial advantages for immobilizing the template polynucleotides and for polymerase extension reaction. First, because PEM can easily be terminated with polymers bearing carboxylic acids, it is easy to attach polynucleotides. Second, the attached template is active for extension by polymerases – most probably, the repulsion of like charges prevents the template from “laying down” on the surface. Finally, the negative charge repels nucleotides, and nonspecific binding is low.

In some other methods of the present invention, only a small percentage of each type of nucleotides present in the extension reaction is labeled, e.g., with fluorescent

dye. As a result, relatively small numbers of incorporated nucleotides are fluorescently labeled, interference of energy transfer is minimized, and the polymerase is less likely to fall off the template or be "choked" by incorporation of two labeled nucleotides sequentially. Optionally, the incorporated fluorescent signals are extinguished by photobleaching.

- 5 Employment of photobleaching strategy can reduce the number of steps (e.g., it may not be necessary to perform the removal of label after every extension cycle). These advantages lead to more accurate detection of incorporated signals, more efficient consumption of polymerase, and a fast sequencing method.

10 II. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which this invention pertains. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY (2d ed. 1994); THE CAMBRIDGE DICTIONARY OF SCIENCE AND TECHNOLOGY (Walker ed., 1988); and Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY (1991). Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are described. The following definitions are provided to assist the reader in the practice of the invention.

The terms "nucleic acid" or "nucleic acid molecule" refer to a deoxyribonucleotide or ribonucleotide polymer in either single- or double-stranded form, and unless otherwise limited, can encompass known analogs of natural nucleotides that can function in a similar manner as naturally occurring nucleotides.

25 "Nucleoside" includes natural nucleosides, including ribonucleosides and 2'-deoxyribonucleosides, as well as nucleoside analogs having modified bases or sugar backbones.

A "base" or "base-type" refers to a particular type of nucleosidic base, such as adenine, cytosine, guanine, thymine, uracil, 5-methylcytosine, 5-bromouracil, 2-aminopurine, deoxyinosine, N<sup>4</sup>-methoxydeoxycytosine, and the like.

30 "Oligonucleotide" or "polynucleotide" refers to a molecule comprised of a plurality of deoxyribonucleotides or nucleoside subunits. The linkage between the nucleoside subunits can be provided by phosphates, phosphonates, phosphoramidates, phosphorothioates, or the like, or by nonphosphate groups as are known in the art, such as

peptoid-type linkages utilized in peptide nucleic acids (PNAs). The linking groups can be chiral or achiral. The oligonucleotides or polynucleotides can range in length from 2 nucleoside subunits to hundreds or thousands of nucleoside subunits. While oligonucleotides are preferably 5 to 100 subunits in length, and more preferably, 5 to 60 subunits in length, the length of polynucleotides can be much greater (e.g., up to 100 kb).

Specific hybridization refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under stringent conditions. Stringent conditions are conditions under which a probe can hybridize to its target subsequence, but to no other sequences. Stringent conditions are sequence-dependent and are different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point ( $T_m$ ) for the specific sequence at a defined ionic strength and pH. The  $T_m$  is the temperature (under defined ionic strength, pH, and nucleic acid concentration) at which 50% of the probes complementary to the target sequence hybridize to the target sequence at equilibrium. Typically, stringent conditions include a salt concentration of at least about 0.01 to 1.0 M Na ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (e.g., 10 to 50 nucleotides). Stringent conditions can also be achieved with the addition of destabilizing agents such as formamide or tetraalkyl ammonium salts. For example, conditions of 5X SSPE (750 mM NaCl, 50 mM Na Phosphate, 5 mM EDTA, pH 7.4) and a temperature of 25-30°C are suitable for allele-specific probe hybridizations. (See Sambrook *et al.*, *Molecular Cloning* 1989).

By "analysis of polynucleotide sequence of a template" is meant determining a sequence of at least 3 contiguous base subunits in a sample fragment, or alternatively, where sequence information is available for a single base-type, the relative positions of at least 3 subunits of identical base-types occurring in sequential order in the fragment. An example of the latter meaning is a determined sequence "AXXAXA" (5'>3'), where a series of 3 adenine (A) bases are found to be separated by two and then one other base-type in the sample fragment.

The term "primer" refers to an oligonucleotide, whether occurring naturally as in a purified restriction digest or produced synthetically, which is capable of acting as a point of initiation of synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, (i.e., in the presence of nucleotides and an inducing agent such as DNA polymerase and at a suitable temperature and pH). The primer is preferably single stranded for maximum efficiency in

amplification, but can alternatively be double stranded. If double stranded, the primer is first treated to separate its strands before being used to prepare extension products. Preferably, the primer is an oligodeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the inducing agent. The exact lengths of the primers depend on many factors, including temperature, source of primer and the use of the method.

A primer is selected to be "substantially" complementary to a strand of specific sequence of the template. A primer must be sufficiently complementary to hybridize with a template strand for primer elongation to occur. A primer sequence need not reflect the exact sequence of the template. For example, a non-complementary nucleotide fragment can be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Non-complementary bases or longer sequences can be interspersed into the primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridize and thereby form a template primer complex for synthesis of the extension product of the primer.

The term "probe" refers to an oligonucleotide (i.e., a sequence of nucleotides), whether occurring naturally as in a purified restriction digest or produced synthetically, recombinantly or by PCR amplification, which is capable of hybridizing to another oligonucleotide of interest. A probe can be single-stranded or double-stranded. Probes are useful in the detection, identification and isolation of particular gene sequences. It is contemplated that any probe used in the present invention can be labeled with any "reporter molecule," so that is detectable in any detection system, including, but not limited to fluorescent, enzyme (e.g., ELISA, as well as enzyme-based histochemical assays), radioactive, quantum dots, and luminescent systems. It is not intended that the present invention be limited to any particular detection system or label.

The term "template," refers to nucleic acid that is to acted upon, such as nucleic acid that is to be mixed with polymerase. In some cases "template" is sought to be sorted out from other nucleic acid sequences. "Substantially single-stranded template" is nucleic acid that is either completely single-stranded (having no double-stranded areas) or single-stranded except for a proportionately small area of double-stranded nucleic acid (such as the area defined by a hybridized primer or the area defined by intramolecular bonding). "Substantially double-stranded template" is nucleic acid that is either completely double-stranded (having no single-stranded region) or double-stranded except for a proportionately small area of single-stranded nucleic acid.

### III. Sequencing Apparatuses

#### A. Basic features of the apparatuses

5 The apparatuses comprise microfabricated channels to which polynucleotide templates to be sequenced are attached. Optionally, the apparatuses comprise plumbing components (e.g., pumps, valves, and connecting channels) for flowing reaction reagents. The apparatuses can also comprise an array of reservoirs for storing reaction reagents (e.g., the polymerase, each type of nucleotides, and other reagents can each be stored in a different reservoir).

10 The microfabricated components of the apparatuses all have a basic "flow channel" structure. The term "flow channel" or "microfabricated flow channel" refers to recess in a structure which can contain a flow of fluid or gas. The polynucleotide templates are attached to the interior surface of microfabricated channels in which synthesis occurs. For consistency and clarity, the flow channels are termed "synthesis channel" when referring to such specific use. The microfabricated flow channels can also be actuated to function as the plumbing components (e.g., micro-pumps, micro-valves, or connecting channels) of the apparatuses.

15 In some applications, microfabricated flow channels are cast on a chip (e.g., a elastomeric chip). Synthesis channels are formed by bonding the chip to a flat substrate (e.g., a glass cover slip) which seals the channel. Thus, one side of the synthesis channel is provided by the flat substrate. Typically, the polynucleotide templates are attached to the interior surface of the substrate within the synthesis channel.

20 The plumbing components can be microfabricated as described in the present invention. For example, the apparatuses can contain in an integrated system a flow cell in which a plurality of synthesis channels are present, and fluidic components (such as micro-pumps, micro-valves, and connecting channels) for controlling the flow of the reagents into and out of the flow cell. Alternatively, the sequencing apparatuses of the present invention utilize plumbing devices described in, e.g., Zdeblick et al., *A Microminiature Electric-to-Fluidic Valve*, Proceedings of the 4th International Conference on Solid State Transducers and Actuators, 1987; Shoji et al., *Smallest Dead Volume Microvalves for Integrated Chemical Analyzing Systems*, Proceedings of Transducers '91, San Francisco, 1991; Vieider et al., *A Pneumatically Actuated Micro Valve with a Silicon Rubber Membrane for Integration with Fluid Handling Systems*, Proceedings of Transducers '95, Stockholm, 1995.

As noted above, at least some of the components of the apparatuses are microfabricated. Employment of microfabricated synthesis channels and/or microfabricated plumbing components significantly reduce the dead volume and decrease the amount of time needed to exchange reagents, which in turn increase the throughput. Microfabrication refers to feature dimensions on the micron level, with at least one dimension of the microfabricated structure being less than 1000  $\mu\text{m}$ . In some apparatuses, only the synthesis channels are microfabricated. In some apparatuses, in addition to the synthesis channels, the valves, pumps, and connecting channels are also microfabricated. Unless otherwise specified, the discussion below of microfabrication is applicable to production of all microfabricated components of the sequencing apparatuses (e.g., the synthesis channels in which sequencing reactions occur, and the valves, pumps, and connecting channels for controlling reagents flow to the synthesis channels).

Various materials can be used to fabricate the microfabricated components (see, e.g., Unger et al., Science 288:113-116, 2000). Preferably, elastomeric materials are used. Thus, in some apparatuses, the integrated (i.e., monolithic) microstructures are made out of various layers of elastomer bonded together. By bonding these various elastomeric layers together, the recesses extending along the various elastomeric layers form flow channels through the resulting monolithic, integral elastomeric structure.

In general, the microfabricated structures (e.g., synthesis channels, pumps, valves, and connecting channels) have widths of about 0.01 to 1000 microns, and a width-to-depth ratios of between 0.1:1 to 100:1. Preferably, the width is in the range of 10 to 200 microns, a width-to-depth ratio of 3:1 to 15:1.

## B. Microfabrication with elastomeric materials

### 1. Basic methods of microfabrication

Various methods can be used to produce the microfabricated components of the sequencing apparatuses of the present invention. Fabrication of the microchannels, valves, pumps can be performed as described in Unger et al., Science 288:113-116, 2000, which is incorporated herein by reference. In some methods (Figs. 1 to 7B, pre-cured elastomer layers are assembled and bonded to produce a flow channel. As illustrated in Fig. 1, a first micro-machined mold 10 is provided. Micro-machined mold 10 can be fabricated by a number of conventional silicon processing methods, including but not limited to photolithography, ion-milling, and electron beam lithography. The micro-machined mold 10 has a raised line or protrusion 11 extending therealong. A first elastomeric layer 20 is cast on

top of mold 10 such that a first recess 21 can be formed in the bottom surface of elastomeric layer 20, (recess 21 corresponding in dimension to protrusion 11), as shown.

As can be seen in Fig. 2, a second micro-machined mold 12 having a raised protrusion 13 extending therealong is also provided. A second elastomeric layer 22 is cast on top of mold 12, as shown, such that a recess 23 can be formed in its bottom surface corresponding to the dimensions of protrusion 13.

As can be seen in the sequential steps illustrated in Figs. 3 and 4, second elastomeric layer 22 is then removed from mold 12 and placed on top of first elastomeric layer 20. As can be seen, recess 23 extending along the bottom surface of second elastomeric layer 22 forms a flow channel 32.

Referring to Fig. 5, the separate first and second elastomeric layers 20 and 22 (Fig. 4) are then bonded together to form an integrated (i.e.: monolithic) elastomeric structure 24.

As can be seen in the sequential step of Figs. 6 and 7A, elastomeric structure 24 is then removed from mold 10 and positioned on top of a planar substrate 14. As can be seen in Fig. 7A and 7B, when elastomeric structure 24 has been sealed at its bottom surface to planar substrate 14, recess 21 forms a flow channel 30.

The present elastomeric structures form a reversible hermetic seal with nearly any smooth planar substrate. An advantage to forming a seal this way is that the elastomeric structures can be peeled up, washed, and re-used. In some apparatuses, planar substrate 14 is glass. A further advantage of using glass is that glass is transparent, allowing optical interrogation of elastomer channels and reservoirs. Alternatively, the elastomeric structure can be bonded onto a flat elastomer layer by the same method as described above, forming a permanent and high-strength bond. This can prove advantageous when higher back pressures are used.

In some methods, microfabrication involves curing each layer of elastomer "in place" (Figs. 8 to 18). In these methods, flow and control channels are defined by first patterning sacrificial layer on the surface of an elastomeric layer (or other substrate, which can include glass) leaving a raised line of sacrificial layer where a channel is desired. Next, a second layer of elastomer is added thereover and a second sacrificial layer is patterned on the second layer of elastomer leaving a raised line of sacrificial layer where a channel is desired. A third layer of elastomer is deposited thereover. Finally, the sacrificial layer is removed by dissolving it out of the elastomer with an appropriate solvent, with the voids formed by removal of the sacrificial layer becoming the flow channels passing through the substrate.

Referring first to Fig. 8, a planar substrate 40 is provided. A first elastomeric layer 42 is then deposited and cured on top of planar substrate 40. Referring to Fig. 9, a first sacrificial layer 44A is then deposited over the top of elastomeric layer 42. Referring to Fig. 10, a portion of sacrificial layer 44A is removed such that only a first line of sacrificial layer 44B remains as shown. Referring to Fig. 11, a second elastomeric layer 46 is then deposited over the top of first elastomeric layer 42 and over the first line of sacrificial layer 44B as shown, thereby encasing first line of sacrificial layer 44B between first elastomeric layer 42 and second elastomeric layer 46. Referring to Fig. 12, elastomeric layers 46 is then cured on layer 42 to bond the layers together to form a monolithic elastomeric substrate 45.

Referring to Fig. 13, a second sacrificial layer 48A is then deposited over elastomeric structure 45. Referring to Fig. 14, a portion of second sacrificial layer 48A is removed, leaving only a second sacrificial layer 48B on top of elastomeric structure 45 as shown. Referring to Fig. 15, a third elastomeric layer 50 is then deposited over the top of elastomeric structure 45 (comprised of second elastomeric layer 42 and first line of sacrificial layer 44B) and second sacrificial layer 48B as shown, thereby encasing the second line of sacrificial layer 48B between elastomeric structure 45 and third elastomeric layer 50.

Referring to Fig. 16, third elastomeric layer 50 and elastomeric structure 45 (comprising first elastomeric layer 42 and second elastomeric layer 46 bonded together) is then bonded together forming a monolithic elastomeric structure 47 having sacrificial layers 44B and 48B passing therethrough as shown. Referring to Fig. 17, sacrificial layers 44B and 48B are then removed (for example, by a solvent ) such that a first flow channel 60 and a second flow channel 62 are provided in their place, passing through elastomeric structure 47 as shown. Lastly, referring to Fig. 18, planar substrate 40 can be removed from the bottom of the integrated monolithic structure.

## 2. Multilayer construction

Soft lithographic bonding can be used to construct an integrated system which contains multiple flow channels. A heterogenous bonding can be used in which different layers are of different chemistries. For example, the bonding process used to bind respective elastomeric layers together can comprise bonding together two layers of RTV 615 silicone. RTV 615 silicone is a two-part addition-cure silicone rubber. Part A contains vinyl groups and catalyst; part B contains silicon hydride (Si-H) groups. The conventional ratio for RTV 615 is 10A:1B. For bonding, one layer can be made with 30A:1B (i.e. excess vinyl groups) and the other with 3A:1B (i.e. excess Si-H groups). Each layer is cured separately. When the



two layers are brought into contact and heated at elevated temperature, they bond irreversibly forming a monolithic elastomeric substrate.

A homogenous bonding can also be used in which all layers are of the same chemistry. For example, elastomeric structures are formed utilizing Sylgard 182, 184 or 186, or aliphatic urethane diacrylates such as (but not limited to) Ebecryl 270 or Irr 245 from UCB Chemical. For example, two-layer elastomeric structures were fabricated from pure acrylated Urethane Ebe 270. A thin bottom layer was spin coated at 8000 rpm for 15 seconds at 170°C. The top and bottom layers were initially cured under ultraviolet light for 10 minutes under nitrogen utilizing a Model ELC 500 device manufactured by Electrolite corporation. The assembled layers were then cured for an additional 30 minutes. Reaction was catalyzed by a 0.5% vol/vol mixture of Irgacure 500 manufactured by Ciba-Geigy Chemicals. The resulting elastomeric material exhibited moderate elasticity and adhesion to glass.

In some applications, two-layer elastomeric structures were fabricated from a combination of 25% Ebe 270 / 50% Irr245 / 25% isopropyl alcohol for a thin bottom layer, and pure acrylated Urethane Ebe 270 as a top layer. The thin bottom layer was initially cured for 5 min, and the top layer initially cured for 10 minutes, under ultraviolet light under nitrogen utilizing a Model ELC 500 device manufactured by Electrolite corporation. The assembled layers were then cured for an additional 30 minutes. Reaction was catalyzed by a 0.5% vol/vol mixture of Irgacure 500 manufactured by Ciba-Geigy Chemicals. The resulting elastomeric material exhibited moderate elasticity and adhered to glass.

Where encapsulation of sacrificial layers is employed to fabricate the elastomer structure as described above in Figs. 8-18, bonding of successive elastomeric layers can be accomplished by pouring uncured elastomer over a previously cured elastomeric layer and any sacrificial material patterned thereupon. Bonding between elastomer layers occurs due to interpenetration and reaction of the polymer chains of an uncured elastomer layer with the polymer chains of a cured elastomer layer. Subsequent curing of the elastomeric layer creates a bond between the elastomeric layers and create a monolithic elastomeric structure.

Referring to the first method of Figs. 1 to 7B, first elastomeric layer 20 can be created by spin-coating an RTV mixture on microfabricated mold 12 at 2000 rpm's for 30 seconds yielding a thickness of approximately 40 microns. Second elastomeric layer 22 can be created by spin-coating an RTV mixture on microfabricated mold 11. Both layers 20 and 22 can be separately baked or cured at about 80°C for 1.5 hours. The second elastomeric layer 22 can be bonded onto first elastomeric layer 20 at about 80°C for about 1.5 hours.

Micromachined molds 10 and 12 can be patterned sacrificial layer on silicon wafers. In an exemplary aspect, a Shipley SJR 5740 sacrificial layer was spun at 2000 rpm patterned with a high resolution transparency film as a mask and then developed yielding an inverse channel of approximately 10 microns in height. When baked at approximately 200°C for about 30 minutes, the sacrificial layer reflows and the inverse channels become rounded. In preferred aspects, the molds can be treated with trimethylchlorosilane (TMCS) vapor for about a minute before each use in order to prevent adhesion of silicone rubber.

### 3. Suitable materials

Allcock et al, Contemporary Polymer Chemistry, 2<sup>nd</sup> Ed. describes elastomers in general as polymers existing at a temperature between their glass transition temperature and liquefaction temperature. Elastomeric materials exhibit elastic properties because the polymer chains readily undergo torsional motion to permit uncoiling of the backbone chains in response to a force, with the backbone chains recoiling to assume the prior shape in the absence of the force. In general, elastomers deform when force is applied, but then return to their original shape when the force is removed. The elasticity exhibited by elastomeric materials can be characterized by a Young's modulus. Elastomeric materials having a Young's modulus of between about 1 Pa – 1 TPa, more preferably between about 10 Pa – 100 GPa, more preferably between about 20 Pa – 1 GPa, more preferably between about 50 Pa – 10 MPa, and more preferably between about 100 Pa – 1 MPa are useful in accordance with the present invention, although elastomeric materials having a Young's modulus outside of these ranges could also be utilized depending upon the needs of a particular application.

The systems of the present invention can be fabricated from a wide variety of elastomers. For example, elastomeric layers 20, 22, 42, 46 and 50 can preferably be fabricated from silicone rubber. In some applications, microstructures of the present systems are fabricated from an elastomeric polymer such as GE RTV 615 (formulation), a vinyl-silane crosslinked (type) silicone elastomer (family). An important requirement for the preferred method of fabrication is the ability to bond multiple layers of elastomers together. In the case of multilayer soft lithography, layers of elastomer are cured separately and then bonded together. This scheme requires that cured layers possess sufficient reactivity to bond together. Either the layers can be of the same type, and are capable of bonding to themselves, or they can be of two different types, and are capable of bonding to each other. Other possibilities include the use an adhesive between layers and the use of thermoset elastomers.

Given the tremendous diversity of polymer chemistries, precursors, synthetic methods, reaction conditions, and potential additives, there are a huge number of possible elastomer systems that could be used to make monolithic elastomeric microstructures.

Variations in the materials used most likely are driven by the need for particular material properties, i.e. solvent resistance, stiffness, gas permeability, or temperature stability.

Common elastomeric polymers include, but are not limited to, polyisoprene, polybutadiene, polychloroprene, polyisobutylene, poly(styrene-butadiene-styrene), the polyurethanes, and silicones. The following is a non-exclusive list of elastomeric materials which can be utilized in connection with the present invention: polyisoprene, polybutadiene, polychloroprene, polyisobutylene, poly(styrene-butadiene-styrene), the polyurethanes, and silicone polymers; or poly(bis(fluoroalkoxy)phosphazene) (PNF, Eypel-F), poly(carborane-siloxanes) (Dexsil), poly(acrylonitrile-butadiene) (nitrile rubber), poly(1-butene), poly(chlorotrifluoroethylene-vinylidene fluoride) copolymers (Kel-F), poly(ethyl vinyl ether), poly(vinylidene fluoride), poly(vinylidene fluoride - hexafluoropropylene) copolymer (Viton), elastomeric compositions of polyvinylchloride (PVC), polysulfone, polycarbonate, polymethylmethacrylate (PMMA), and polytetrafluoroethylene (Teflon).

In addition, polymers incorporating materials such as chlorosilanes or methyl-, ethyl-, and phenylsilanes, and polydimethylsiloxane (PDMS) such as Dow Chemical Corp. Sylgard 182, 184 or 186, or aliphatic urethane diacrylates such as (but not limited to) Ebecryl 270 or Irr 245 from UCB Chemical can also be used.

In some methods, elastomers can also be "doped" with uncrosslinkable polymer chains of the same class. For instance RTV 615 can be diluted with GE SF96-50 Silicone Fluid. This serves to reduce the viscosity of the uncured elastomer and reduces the Young's modulus of the cured elastomer. Essentially, the crosslink-capable polymer chains are spread further apart by the addition of "inert" polymer chains, so this is called "dilution". RTV 615 cures at up to 90% dilution, with a dramatic reduction in Young's modulus.

Other examples of doping of elastomer material can include the introduction of electrically conducting or magnetic species. Should it be desired, doping with fine particles of material having an index of refraction different than the elastomeric material (i.e. silica, diamond, sapphire) is also contemplated as a system for altering the refractive index of the material. Strongly absorbing or opaque particles can be added to render the elastomer colored or opaque to incident radiation. This can conceivably be beneficial in an optically addressable system.

4. Dimensions of the microfabricated structures

Some flow channels (30, 32, 60 and 62) preferably have width-to-depth ratios of about 10:1. A non-exclusive list of other ranges of width-to-depth ratios in accordance with the present invention is 0.1:1 to 100:1, more preferably 1:1 to 50:1, more preferably 2:1 to 20:1, and most preferably 3:1 to 15:1. In an exemplary aspect, flow channels 30, 32, 60 and 62 have widths of about 1 to 1000 microns. A non-exclusive list of other ranges of widths of flow channels in accordance with the present invention is 0.01 to 1000 microns, more preferably 0.05 to 1000 microns, more preferably 0.2 to 500 microns, more preferably 1 to 250 microns, and most preferably 10 to 200 microns. Exemplary channel widths include 0.1  $\mu\text{m}$ , 1  $\mu\text{m}$ , 2  $\mu\text{m}$ , 5  $\mu\text{m}$ , 10  $\mu\text{m}$ , 20  $\mu\text{m}$ , 30  $\mu\text{m}$ , 40  $\mu\text{m}$ , 50  $\mu\text{m}$ , 60  $\mu\text{m}$ , 70  $\mu\text{m}$ , 80  $\mu\text{m}$ , 90  $\mu\text{m}$ , 100  $\mu\text{m}$ , 110  $\mu\text{m}$ , 120  $\mu\text{m}$ , 130  $\mu\text{m}$ , 140  $\mu\text{m}$ , 150  $\mu\text{m}$ , 160  $\mu\text{m}$ , 170  $\mu\text{m}$ , 180  $\mu\text{m}$ , 190  $\mu\text{m}$ , 200  $\mu\text{m}$ , 210  $\mu\text{m}$ , 220  $\mu\text{m}$ , 230  $\mu\text{m}$ , 240  $\mu\text{m}$ , and 250  $\mu\text{m}$ .

Flow channels 30, 32, 60, and 62 have depths of about 1 to 100 microns. A non-exclusive list of other ranges of depths of flow channels in accordance with the present invention is 0.01 to 1000 microns, more preferably 0.05 to 500 microns, more preferably 0.2 to 250 microns, and more preferably 1 to 100 microns, more preferably 2 to 20 microns, and most preferably 5 to 10 microns. Exemplary channel depths include including 0.01  $\mu\text{m}$ , 0.02  $\mu\text{m}$ , 0.05  $\mu\text{m}$ , 0.1  $\mu\text{m}$ , 0.2  $\mu\text{m}$ , 0.5  $\mu\text{m}$ , 1  $\mu\text{m}$ , 2  $\mu\text{m}$ , 3  $\mu\text{m}$ , 4  $\mu\text{m}$ , 5  $\mu\text{m}$ , 7.5  $\mu\text{m}$ , 10  $\mu\text{m}$ , 12.5  $\mu\text{m}$ , 15  $\mu\text{m}$ , 17.5  $\mu\text{m}$ , 20  $\mu\text{m}$ , 22.5  $\mu\text{m}$ , 25  $\mu\text{m}$ , 30  $\mu\text{m}$ , 40  $\mu\text{m}$ , 50  $\mu\text{m}$ , 75  $\mu\text{m}$ , 100  $\mu\text{m}$ , 150  $\mu\text{m}$ , 200  $\mu\text{m}$ , and 250  $\mu\text{m}$ .

The flow channels are not limited to these specific dimension ranges and examples given above, and can vary in width in order to affect the magnitude of force required to deflect the membrane as discussed at length below in conjunction with Fig. 21. For example, extremely narrow flow channels having a width on the order of 0.01  $\mu\text{m}$  can be useful in optical and other applications, as discussed in detail below. Elastomeric structures which include portions having channels of even greater width than described above are also contemplated by the present invention, and examples of applications of utilizing such wider flow channels include fluid reservoir and mixing channel structures.

Elastomeric layer 22 can be cast thick for mechanical stability. In an exemplary embodiment, layer 22 is 50 microns to several centimeters thick, and more preferably approximately 4 mm thick. A non-exclusive list of ranges of thickness of the elastomer layer in accordance with other embodiments of the present invention is between about 0.1 micron to 10 cm, 1 micron to 5 cm, 10 microns to 2 cm, 100 microns to 10 mm.

Accordingly, membrane 25 of Fig. 7B separating flow channels 30 and 32 has a typical thickness of between about 0.01 and 1000 microns, more preferably 0.05 to 500 microns, more preferably 0.2 to 250, more preferably 1 to 100 microns, more preferably 2 to 50 microns, and most preferably 5 to 40 microns. As such, the thickness of elastomeric layer 22 is about 100 times the thickness of elastomeric layer 20. Exemplary membrane thicknesses include 0.01  $\mu\text{m}$ , 0.02  $\mu\text{m}$ , 0.03  $\mu\text{m}$ , 0.05  $\mu\text{m}$ , 0.1  $\mu\text{m}$ , 0.2  $\mu\text{m}$ , 0.3  $\mu\text{m}$ , 0.5  $\mu\text{m}$ , 1  $\mu\text{m}$ , 2  $\mu\text{m}$ , 3  $\mu\text{m}$ , 5  $\mu\text{m}$ , 7.5  $\mu\text{m}$ , 10  $\mu\text{m}$ , 12.5  $\mu\text{m}$ , 15  $\mu\text{m}$ , 17.5  $\mu\text{m}$ , 20  $\mu\text{m}$ , 22.5  $\mu\text{m}$ , 25  $\mu\text{m}$ , 30  $\mu\text{m}$ , 40  $\mu\text{m}$ , 50  $\mu\text{m}$ , 75  $\mu\text{m}$ , 100  $\mu\text{m}$ , 150  $\mu\text{m}$ , 200  $\mu\text{m}$ , 250  $\mu\text{m}$ , 300  $\mu\text{m}$ , 400  $\mu\text{m}$ , 500  $\mu\text{m}$ , 750  $\mu\text{m}$ , and 1000  $\mu\text{m}$

Similarly, first elastomeric layer 42 can have a preferred thickness about equal to that of elastomeric layer 20 or 22; second elastomeric layer 46 can have a preferred thickness about equal to that of elastomeric layer 20; and third elastomeric layer 50 can have a preferred thickness about equal to that of elastomeric layer 22.

#### C. Operation of the microfabricated components

Figs. 7B and 7C together show the closing of a first flow channel by pressurizing a second flow channel, with Fig. 7B (a front sectional view cutting through flow channel 32 in corresponding Fig. 7A), showing an open first flow channel 30; with Fig. 7C showing first flow channel 30 closed by pressurization of the second flow channel 32.

Referring to Fig. 7B, first flow channel 30 and second flow channel 32 are shown. Membrane 25 separates the flow channels, forming the top of first flow channel 30 and the bottom of second flow channel 32. As can be seen, flow channel 30 is "open".

As can be seen in Fig. 7C, pressurization of flow channel 32 (either by gas or liquid introduced therein) causes membrane 25 to deflect downward, thereby pinching off flow F passing through flow channel 30. Accordingly, by varying the pressure in channel 32, a linearly actuatable valving system is provided such that flow channel 30 can be opened or closed by moving membrane 25 as desired.

It is to be understood that exactly the same valve opening and closing methods can be achieved with flow channels 60 and 62. Since such valves are actuated by moving the roof of the channels themselves (i.e., moving membrane 25), valves and pumps produced by this technique have a truly zero dead volume, and switching valves made by this technique have a dead volume approximately equal to the active volume of the valve, for example about  $100 \times 100 \times 10 \mu\text{m} = 100\text{pL}$ . Such dead volumes and areas consumed by the moving

membrane are approximately two orders of magnitude smaller than known conventional microvalves. Smaller and larger valves and switching valves are contemplated in the present invention, and a non-exclusive list of ranges of dead volume includes 1 aL to 1 uL, 100 aL to 100 nL, 1 fL to 10 nL, 100 fL to 1 nL, and 1 pL to 100 pL

5 The extremely small volumes capable of being delivered by pumps and valves in accordance with the present invention represent a substantial advantage. Specifically, the smallest known volumes of fluid capable of being manually metered is around 0.1  $\mu$ l. The smallest known volumes capable of being metered by automated systems is about ten-times larger (1  $\mu$ l). Utilizing pumps and valves of the present invention, volumes of liquid of 10 nl  
10 or smaller can routinely be metered and dispensed. The accurate metering of extremely small volumes of fluid enabled by the present invention would be extremely valuable in a large number of biological applications, including diagnostic tests and assays.

Figs. 21a and 21b illustrate valve opening vs. applied pressure for a 100  $\mu$ m wide first flow channel 30 and a 50  $\mu$ m wide second flow channel 32. The membrane of this  
15 device was formed by a layer of General Electric Silicones RTV 615 having a thickness of approximately 30 $\mu$ m and a Young's modulus of approximately 750 kPa. Figs. 21a and 21b show the extent of opening of the valve to be substantially linear over most of the range of applied pressures.

Air pressure was applied to actuate the membrane of the device through a 10  
20 cm long piece of plastic tubing having an outer diameter of 0.025" connected to a 25 mm piece of stainless steel hypodermic tubing with an outer diameter of 0.025" and an inner diameter of 0.013". This tubing was placed into contact with the control channel by insertion into the elastomeric block in a direction normal to the control channel. Air pressure was applied to the hypodermic tubing from an external LHDA miniature solenoid valve  
25 manufactured by Lee Co.

The response of valves of the present invention is almost perfectly linear over a large portion of its range of travel, with minimal hysteresis. While valves and pumps do not require linear actuation to open and close, linear response does allow valves to more easily be used as metering devices. In some applications, the opening of the valve is used to  
30 control flow rate by being partially actuated to a known degree of closure. Linear valve actuation makes it easier to determine the amount of actuation force required to close the valve to a desired degree of closure. Another benefit of linear actuation is that the force required for valve actuation can be easily determined from the pressure in the flow channel.

If actuation is linear, increased pressure in the flow channel can be countered by adding the same pressure (force per unit area) to the actuated portion of the valve.

D. Schematic illustration of the elastomeric apparatuses

5 An exemplary sequencing system is illustrated in Fig. 25. Four reservoirs 150A, 150B, 150C and 150D have labeled nucleotides A, C, T and G respectively disposed therein. Four flow channels 30A, 30B, 30C and 30D are connected to reservoirs 150A, 150B, 150C and 150D. Four control lines 32A, 32B, 32C and 32D (shown in phantom) are disposed thereacross with control line 32A permitting flow only through flow channel 30A  
10 (i.e.: sealing flow channels 30B, 30C and 30D), when control line 32A is pressurized. Similarly, control line 32B permits flow only through flow channel 30B when pressurized. As such, the selective pressurization of control lines 32A, 32B, 32C and 32D sequentially selects a desired nucleotide (A, C, T or G) from a desired reservoir (150A, 150B, 150C or 150D). The fluid then passes through flow channel 120 into a multiplexed channel flow  
15 controller 125, which in turn directs fluid flow into one or more of a plurality of synthesis channels or reaction chambers 122A, 122B, 122C, 122D or 122E in which solid phase synthesis can be carried out.

Fig. 26 illustrates a further extension of the system shown in Fig. 25. It has a plurality of reservoirs R1 to R13. These reservoirs can contain the labeled nucleotides,  
20 nucleotide polymerase, or reagents for coating the surface of the synthesis channel and attaching polynucleotide templates (see below for further discussion). The reservoirs are connected to systems 200 as set forth in Figs. 25. Systems 200 are connected to a multiplexed channel flow controller 125, which is in turn connected to a plurality of synthesis channels or reaction chambers. An advantage of this system is that both of multiplexed  
25 channel flow controllers 125 and fluid selection systems 200 can be controlled by the same pressure inputs 170 and 172, provided a "close horizontal" and a "close vertical" control lines (160 and 162, in phantom) are also provided.

Some apparatuses comprise a plurality of selectively addressable reaction chambers that are disposed along a flow channel. In these apparatuses, the polynucleotide  
30 templates can be attached to the surface of the reaction chambers instead of the surface of flow channels. An exemplary embodiment of such apparatuses is illustrated in Figs. 22A, 22B, 22C and 22D. It is a system for selectively directing fluid flow into one or more of a plurality of reaction chambers disposed along a flow channel.

Fig. 22A shows a top view of a flow channel 30 having a plurality of reaction chambers 80A and 80B disposed therealong. Preferably flow channel 30 and reaction chambers 80A and 80B are formed together as recesses into the bottom surface of a first layer 100 of elastomer.

5 Fig. 22B shows a bottom plan view of another elastomeric layer 110 with two control lines 32A and 32B each being generally narrow, but having wide extending portions 33A and 33B formed as recesses therein.

As seen in the exploded view of Fig. 22C, and assembled view of Fig. 22D, elastomeric layer 110 is placed over elastomeric layer 100. Layers 100 and 110 are then  
10 bonded together, and the integrated system operates to selectively direct fluid flow F (through flow channel 30) into either or both of reaction chambers 80A and 80B, as follows.

Pressurization of control line 32A will cause the membrane 25 (i.e.: the thin portion of elastomer layer 100 located below extending portion 33A and over regions 82A of reaction chamber 80A) to become depressed, thereby shutting off fluid flow passage in regions 82A,  
15 effectively sealing reaction chamber 80 from flow channel 30. As can also be seen, extending portion 33A is wider than the remainder of control line 32A. As such, pressurization of control line 32A will not result in control line 32A sealing flow channel 30.

As can be appreciated, either or both of control lines 32A and 32B can be actuated at once. When both control lines 32A and 32B are pressurized together, sample flow  
20 in flow channel 30 will enter neither of reaction chambers 80A or 80B.

The concept of selectably controlling fluid introduction into various addressable reaction chambers disposed along a flow line (Figs. 22) can be combined with concept of selectably controlling fluid flow through one or more of a plurality of parallel flow lines (Fig. 21) to yield a system in which a fluid sample or samples can be sent to any  
25 particular reaction chamber in an array of reaction chambers. An example of such a system is provided in Fig. 23, in which parallel control channels 32A, 32B and 32C with extending portions 34 (all shown in phantom) selectively direct fluid flows F1 and F2 into any of the array of reaction wells 80A, 80B, 80C or 80D as explained above; while pressurization of control lines 32C and 32D selectively shuts off flows F2 and F1, respectively.

30 In yet another embodiment, fluid passage between parallel flow channels is possible. Referring to Fig. 24, either or both of control lines 32A or 32D can be depressurized such that fluid flow through lateral passageways 35 (between parallel flow channels 30A and 30B) is permitted. In this aspect of the invention, pressurization of control lines 32C and 32D would shut flow channel 30A between 35A and 35B, and would also shut lateral



passageways 35B. As such, flow entering as flow F1 would sequentially travel through 30A, 35A and leave 30B as flow F4.

E. Non-elastomer based apparatuses

As discussed above, while elastomers are preferred materials for fabricating the sequencing apparatuses of the present invention, non-elastomer based microfluidic devices can also be used in the apparatuses of the present invention. In some applications, the sequencing apparatuses utilize microfluidics based on conventional micro-electro-mechanical system (MEMS) technology. Methods of producing conventional MEMS microfluidic systems such as bulk micro-machining and surface micro-machining have been described, e.g., in Terry et al., *A Gas Chromatographic Air Analyzer Fabricated on a Silicon Wafer*, IEEE Trans. on Electron Devices, v. ED-26, pp. 1880-1886, 1979; and Berg et al., *Micro Total Analysis Systems*, New York, Kluwer, 1994.

Bulk micro-machining is a subtractive fabrication method whereby single crystal silicon is lithographically patterned and then etched to form three-dimensional structures. For example, bulk micromachining technology, which includes the use of glass wafer processing, silicon-to-glass wafer bonding, has been commonly used to fabricate individual microfluidic components. This glass-bonding technology has also been used to fabricate microfluidic systems.

Surface micro-machining is an additive method where layers of semiconductor-type materials such as polysilicon, silicon nitride, silicon dioxide, and various metals are sequentially added and patterned to make three-dimensional structures. Surface micromachining technology can be used to fabricate individual fluidic components as well as microfluidic systems with on-chip electronics. In addition, unlike bonded-type devices, hermetic channels can be built in a relatively simple manner using channel walls made of polysilicon (see, e.g., Webster et al., *Monolithic Capillary Gel Electrophoresis Stage with On-Chip Detector*, in International Conference on Micro Electromechanical Systems, MEMS 96, pp. 491-496, 1996), silicon nitride (see, e.g., Mastrangelo et al., *Vacuum-Sealed Silicon Micromachined Incandescent Light Source*, in Intl. Electron Devices Meeting, IDEM 89, pp. 503-506, 1989), and silicon dioxide.

In some applications, electrokinetic flow based microfluidics can be employed in the sequencing apparatuses of the present invention. Briefly, these systems direct reagents flow within an interconnected channel and/or chamber containing structure through the application of electrical fields to the reagents. The electrokinetic systems concomitantly

regulate voltage gradients applied across at least two intersecting channels. Such systems are described, e.g., in WO 96/04547 and U.S. Patent No. 6,107,044.

An exemplary electrokinetic flow based microfluidic device can have a body structure which includes at least two intersecting channels or fluid conduits, e.g., interconnected, enclosed chambers, which channels include at least three unintersected termini. The intersection of two channels refers to a point at which two or more channels are in fluid communication with each other, and encompasses "T" intersections, cross intersections, "wagon wheel" intersections of multiple channels, or any other channel geometry where two or more channels are in such fluid communication. An unintersected terminus of a channel is a point at which a channel terminates not as a result of that channel's intersection with another channel, e.g., a "T" intersection.

In some electrokinetic flow based apparatuses, at least three intersecting channels having at least four unintersected termini are present. In a basic cross channel structure, where a single horizontal channel is intersected and crossed by a single vertical channel, controlled electrokinetic transport operates to direct reagent flow through the intersection, by providing constraining flows from the other channels at the intersection. Simple electrokinetic flow of this reagent across the intersection could be accomplished by applying a voltage gradient across the length of the horizontal channel, i.e., applying a first voltage to the left terminus of this channel, and a second, lower voltage to the right terminus of this channel, or by allowing the right terminus to float (applying no voltage).

In some other applications, the apparatus comprises a microfabricated flow cell with external mini-fluidics. Such an apparatus is illustrated in Fig. 27. The glass cover slip can be anodically bonded to the surface of the flow cell. The interrogation region is  $100\mu\text{m} \times 100\mu\text{m} \times 100\mu\text{m}$ , while the input and output channels are  $100\mu\text{m} \times 100\mu\text{m} \times 100\mu\text{m}$ . Holes for the attachment of plumbing are etched at the ends of the channels. For such apparatuses, the fluidics can be external. Plumbing can be performed with standard HPLC components, e.g., from Upchurch and Hamilton. In the interrogation region, the polynucleotide template is attached to the surface with standard avidin-biotin chemistry. Multiple copies of templates can be attached to the apparatus. For example, for a 7 kb template, the radius of gyration is approximately  $0.2\mu\text{m}$ . Therefore, about  $10^5$  molecules can be attached while preventing the molecules from touching. Reagent switching can be accomplished with, e.g., an Upchurch six port injection valve and driven by, e.g., a Thar Designs motor. Fluid can be pumped with a syringe pump. The detection system can be an

external optical microscope, with the objective being in close proximity to the glass cover slip.

#### IV Analysis of polynucleotide sequences

##### 5 A. Template preparation and attachment to surface of synthesis channel

###### 1. The general scheme

In some applications, the polynucleotides to be analyzed are first cloned in single-stranded M13 plasmid (see, e.g., Current Protocols In Molecular Biology, Ausubel, et al., eds., John Wiley & Sons, Inc. 1995; and Sambrook, et al., Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Press, 1989). The single stranded plasmid is primed by 5'-biotinylated primers (see, e.g., U.S. Patent No. 5,484,701), and double stranded plasmid can then be synthesized. The double stranded circle is then linearized, and the biotinylated strand is purified. In some methods, templates of around 100 bp in length are analyzed. In some methods, templates to be sequenced are about 1 kb in length. In other methods, templates that can be analyzed have a length of about 3 kb, 10 kb, or 20 kb.

Primer annealing is performed under conditions which are stringent enough to achieve sequence specificity yet sufficiently permissive to allow formation of stable hybrids at an acceptable rate. The temperature and length of time required for primer annealing depend upon several factors including the base composition, length and concentration of the primer, and the nature of the solvent used, e.g., the concentration of DMSO, formamide, or glycerol, and counter ions such as magnesium. Typically, hybridization with synthetic polynucleotides is carried out at a temperature that is approximately 5 to 10°C below the melting temperature of the target-primer hybrid in the annealing solvent. Preferably, the annealing temperature is in the range of 55 to 75°C. and the primer concentration is approximately 0.2 μM. Under these preferred conditions, the annealing reaction can be complete in only a few seconds.

The single stranded polynucleotide templates to be analyzed can be DNA or RNA. They can comprise naturally occurring and or non-naturally occurring nucleotides. Templates suitable for analysis according to the present invention can have various sizes. For example, the template can have a length of 100 bp, 200 bp, 500 bp, 1 kb, 3 kb, 10 kb, or 20 kb.

In some methods, the templates are immobilized to the surface of the synthesis channels (e.g., 122A-122E in Fig. 25). By immobilizing the templates, unincorporated nucleotides can be removed from the synthesis channels by a washing step. The templates

can be immobilized to the surface prior to hybridization to the primer. The templates can also be hybridized to the primers first and then immobilize to the surface. Alternatively, the primers are immobilized to the surface, and the templates are attached to the synthesis channels through hybridization to the primers.

5                Various methods can be used to immobilize the templates or the primers to the surface of the synthesis channels or reaction chambers. The immobilization can be achieved through direct or indirect bonding of the templates to the surface. The bonding can be by covalent linkage. See, Joos et al., *Analytical Biochemistry* 247:96-101, 1997; Oroskar et al., *Clin. Chem* 42:1547-1555, 1996; and Khandjian, *Mole. Bio. Rep.* 11:107-115, 1986. The  
10           bonding can also be through non-covalent linkage. For example, Biotin-streptavidin (Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991) and digoxigenin and anti-digoxigenin (Smith et al., *Science* 253: 1122, 1992) are common tools for attaching polynucleotides to surfaces and parallels. Alternatively, the bonding can be achieved by anchoring a hydrophobic chain into a lipidic monolayer or bilayer.

15                When biotin-streptavidin linkage is used to immobilize the templates, the templates are biotinylated, and one surface of the synthesis channels are coated with streptavidin. Since streptavidin is a tetramer, it has four biotin binding sites per molecule. Thus, in order to coat a surface with streptavidin, the surface can be biotinylated first, and then one of the four binding sites of streptavidin can be used to anchor the protein to the  
20           surface, leaving the other sites free to bind the biotinylated template (see, Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991). Such treatment leads to a high density of streptavidin on the surface of the synthesis channel, allowing a correspondingly high density of template coverage. Reagents for biotinylating a surface can be obtained, for example, from Vector laboratories.

25                In some applications, the substrate or synthesis channel is pretreated to create surface chemistry that facilitates attachment of the polynucleotide templates and subsequent synthesis reactions. In some methods, the surface is coated with a polyelectrolyte multilayer (PEM). Attachment of templates to PEM-coated surface can be accomplished by light-directed spatial attachment (see, e.g., U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837).  
30           Alternatively, the templates can be attached to PEM-coated surface entire chemically (see below for detail). In some methods, non-PEM based surface chemistry can be created prior to template attachment.

## 2. Attachment of diverse templates to a single channel

While diverse polynucleotide templates can be each immobilized to and sequenced in a separate synthesis channel, multiple templates can also be sequenced in a single microfluidic synthesis channel. In the latter scenario, the templates are attached at different locations along the flow path of the channel. This can be accomplished by a variety of different methods, including hybridization of primer capture sequences to oligonucleotides immobilized at different points on the substrate, and sequential activation of different points down the channel towards template immobilization.

Methods of creation of surfaces with arrays of oligonucleotides have been described, e.g., in U.S. Patent Nos. 5,744,305, 5,837,832, and 6,077,674. Such a surface can be used as a substrate that is to be bond to a microfluidic chip and form the synthesis channel. Primers with two domains, a priming domain and a capture domain, can be used to anchor templates to the substrate. The priming domain is complementary to the target template. The capture domain is present on the non-extended side of the priming sequence. It is not complementary to the target template, but rather to a specific oligonucleotide sequence present on the substrate. The target templates can be separately hybridized with their primers, or (if the priming sequences are different) simultaneously hybridized in the same solution. Incubation of the primer/template duplexes in the flow channel under hybridization conditions allows attachment of each template to a unique spot. Multiple synthesis channels can be charged with templates in this fashion simultaneously.

Another method for attaching multiple templates in a single channel is to sequentially activate portions of the substrate and attach template to them. Activation of the substrate can be achieved by either optical or electrical means. Optical illumination can be used to initiate a photochemical deprotection reaction that allows attachment of the template to the surface (see, e.g., U.S. Patent Nos. 5,599,695, 5,831,070, and 5,959,837). For instance, the substrate surface can be derivatized with "caged biotin", a commercially available derivative of biotin that becomes capable of binding to avidin only after being exposed to light. Templates can then be attached by exposure of a site to light, filling the channel with avidin solution, washing, and then flowing biotinylated template into the channel. Another variation is to prepare avidinylated substrate and a template with a primer with a caged biotin moiety; the template can then be immobilized by flowing into the channel and illumination of the solution above a desired area. Activated template/primer duplexes are then attached to the first wall they diffused to, yielding a diffusion limited spot.

Electrical means can also be used to direct template to specific points in the channel. By positively charging one electrode in the channel and negatively charging the

others, a field gradient can be created which drives the template to a single electrode, where it can attach (see, e.g., U.S. Patent Nos. 5,632,957, 6,051,380, and 6,071,394). Alternatively, it can be achieved by electrochemically activating regions of the surface and changing the voltage applied to the electrodes.

5

B. Exemplary surface chemistry for attaching templates: PEM coating

In some methods, the surface of synthesis channels are coated with PEM prior to attachment of the templates (or primers). Such attachment scheme can be both an ex-situ process or an in situ process. With the ex-situ protocol, the surface of the flat substrate is coated with PEM first, followed by attachment of the templates. The elastomeric microfluidic chip is then bonded to the substrate to form and seal the synthesis channel. With the in-situ protocol, the microfluidic chip is attached to the flat substrate first, and a PEM is then constructed in the channels. The templates are then attached inside the channels. In still some other applications, the microfluidic chip can be bonded to the flat substrate at any point in the template attachment process, and the remaining steps can be completed inside the microfluidic channels.

Preferably, the in-situ protocol is used. The method described here leads to low nonspecific binding of labeled (e.g., with fluorescent dye) nucleotides and good seal of the microfluidic components and the synthesis channels. A good seal between the microfluidic components and the synthesis channels allows the use of higher pressures, which in turn increases flow rates and decreases exchange times. The various methods for attaching the templates to the surface of the synthesis channel are discussed in detail below.

An exemplified scheme of the ex situ protocol is as follows. First, the surface of a glass cover slip is cleaned and then coated with a polyelectrolyte multilayer (PEM). Following biotinylation of the carboxylic acid groups, streptavidin is then applied to generate a surface capable of capturing biotinylated molecules. Biotinylated polynucleotide templates are then added to the coated glass cover slip for attachment. The surface chemistry thus created is particularly suited for sequencing by synthesis with fluorescent nucleotides, because it generates a strong negatively-charged surface which repels the negatively-charged nucleotides. Detailed procedures for cleaning the cover slips, coating of polyelectrolyte multilayer, and attachment of the templates are described below.

PEM formation proceeds by the sequential addition of polycations and polyanions, which are polymers with many positive or negative charges, respectively. Upon addition of a polycation to a negatively-charged surface, the polycation deposits on the

surface, forming a thin polymer layer and reversing the surface charge. Similarly, a polyanion deposited on a positively charged surface forms a thin layer of polymer and leaves a negatively charged surface. Alternating exposure to poly(+) and poly(-) generates a polyelectrolyte multilayer structure with a surface charge determined by the last  
5 polyelectrolyte added; in the case of incompletely-charged surfaces, multiple-layer deposition also tends to increase surface charge to a well defined and stable level. PEM formation has been described by Decher et al.(Thin Solid Films, 210:831-835, 1992).

Carboxylic acid groups are negatively charged at pH 7, and are a common target for covalent bond formation. By terminating the surface with carboxylic acid groups, a  
10 surface which is both strongly negatively-charged and chemically reactive can be generated. In particular, amines can link to them to form amide bonds, a reaction that can be catalyzed by carbodiimides. A molecule with biotin at one end, a hydrophilic spacer, and an amine at the other end is used to terminate the surface with biotin.

An avidin molecule is capable of binding up to four biotin molecules. This  
15 means that avidin, and its derivative Streptavidin, is capable of converting a biotin-terminated surface to a surface capable of *capturing* biotin. Streptavidin, which carries a slight negative charge, is used to attach the polynucleotide templates to be analyzed to the surface by using a biotinylated primer. A buffer with a high concentration of multivalent salt is used in order to screen the repulsion of the negatively charged surface for the negatively-charged DNA.

To coat the polyelectrolyte multilayer, the glass cover slips are first cleaned  
20 with HP H<sub>2</sub>O (H<sub>2</sub>O deionized to 18.3 MOhm-cm and filtered to 0.2  $\mu$ m) and a RCA Solution (6:4:1 mixture of HP H<sub>2</sub>O, (30% NH<sub>4</sub>OH), and (30% H<sub>2</sub>O<sub>2</sub>)). The cover slips are then sonicated in 2% Micro 90 detergent for 20 minutes. After rinse thoroughly with HP H<sub>2</sub>O, the cover slips are stirred in gently boiling RCA solution for at least 1 hour, and rinsed again  
25 with HP H<sub>2</sub>O.

After cleaning, the glass cover slips are submerged in PAll solution (Poly(allylamine) (PAll, +): 2 mg/ml in HP H<sub>2</sub>O, adjusted to pH 7.0) and agitate for at least 10 minutes. The cover slips are then removed from PAll and washed with HP H<sub>2</sub>O by submerging in HP H<sub>2</sub>O with agitation for at least three times. The treatment continues by  
30 agitation in a PAcr solution (Poly(acrylic acid) (PAcr, -): 2 mg/ml in HP H<sub>2</sub>O, adjusted to pH 7.0) for at least 10 minutes and washing with HP H<sub>2</sub>O. The treatment steps are then repeated once.

After PEM coating, the PEM coated glass is incubated with a EDC/BLCPA solution for 30 minutes. The EDC/BLCPA solution is prepared by mixing equal amounts of



50 mM EDC solution (in MES buffer) and 50 mM BLCPA (in MES buffer) and diluting to 5mM in MES buffer. The glass is then rinsed with 10 mM Tris-NaCl and incubated with 0.1 mg/ml streptavidin solution for 1 hour. After washing with 10 mM Tris-NaCl, the glass is incubated with a solution containing the polynucleotide template ( $10^{-7}$  M in Tris 100 mM MgCl<sub>2</sub>) for 30 minutes. The glass is again rinsed thoroughly with 10 mM Tris-NaCl.

For in-situ attachment, the microfluidic substrate is bonded to the glass cover slip by HCl-assisted bonding. Essentially, the chips are first washed with a surfactant (e.g., first with HP H<sub>2</sub>O, then in 0.1% Tween 20, then rinse again with HP H<sub>2</sub>O). The washed microfluidic chips are then put on the glass cover slips with a few microliters of dilute HCl (e.g., 1% HCl in HP H<sub>2</sub>O), followed by baking at 37° C for 1-2 hours. Such treatment enhances the bond strength to glass (e.g., >20 psi pressure) without increasing nonspecific adsorption.

Following HCl treatment, PEM formation, biotinylation, streptavidinylation, and template attachment can be performed using essentially the same reagents and methods as described above for ex-situ attachment, except the solutions are injected through the channels by pressure instead of just being aliquoted onto the substrate surface.

Coating the microchannel surface with the PEM technique is significant for analyzing polynucleotide sequences according to the present invention. In general, the method used to attach the template to the surface should fulfill several requirements in order to be useful in a sequencing-by-synthesis application. First, it must be possible to attach reasonable quantities of polynucleotide templates. In addition, the attached templates should remain active for polymerase action. Further, nonspecific binding of fluorescent nucleotides should be very low.

If insufficient numbers of template molecules are bound, the signal-to-noise ratio of the technique is too low to allow useful sequencing. Binding large quantities of templates is insufficient, however, if the primer/target duplex cannot be extended by a polymerase. This is a problem for surface chemistry based on building off amine-bearing surfaces: amines are positively charged at normal pH. This means that the negatively-charged DNA backbone can non-specifically stick to the surface, and that the polymerase is sterically impeded from adding nucleotides. Finally, if there is significant nonspecific binding of fluorescent nucleotides to the surface, it becomes impossible to distinguish between signal due to incorporation and signal due to nonspecific binding.

When the nucleotides are fluorescently labeled, they generally have relatively strong nonspecific binding to many surfaces because they possess both a strongly polar



moiety (the nucleotide, and in particular the triphosphate) and a relatively hydrophobic moiety (the fluorescent dye). A surface bearing positively-charged groups (i.e. amines) invariably has a very high nonspecific binding due to the attraction of the triphosphate group (which is strongly negatively charged) to the positively-charged amines. Neutral surfaces generally have strong nonspecific binding due to the action of the fluorescent nucleotide as a surfactant (i.e. assembling with nonpolar moiety towards the uncharged (more hydrophobic) surface and polar end in the aqueous phase). A surface bearing negative charges can repel the negatively charged fluorescent nucleotides, so it has the lowest nonspecific binding. Glass is such a surface, but the surface silanols that give it its negative charge in water are a difficult target to attach DNA to directly. Typical DNA attachment protocols use silanization (often with aminosilanes) to attach template; as discussed earlier amino groups lead to unacceptable levels of nonspecific binding.

A polyelectrolyte multilayer terminated with carboxylic acid-bearing polymer fulfills all three criteria. First, it is easy to attach polynucleotide to because carboxylic acids are good targets for covalent bond formation. Second, the attached template is active for extension by polymerases – most probably, the repulsion of like charges prevents the template from “laying down” on the surface. Finally, the negative charge repels the fluorescent nucleotides, and nonspecific binding is low.

The attachment scheme described here is easy to generalize on. Without modification, the PEM/biotin/streptavidin surface that is produced can be used to capture or immobilize any biotinylated molecule. A slight modification can be the use of another capture pair, i.e. substituting digoxigenin (dig) for biotin and labeling the molecule to be immobilized with anti-digoxigenin (anti-dig). Reagents for biotinylation or dig-labeling of amines are all commercially available.

Another generalization is that the chemistry is nearly independent of the surface chemistry of the support. Glass, for instance, can support PEMs terminated with either positive or negative polymer, and a wide variety of chemistry for either. But other substrates such as silicone, polystyrene, polycarbonate, etc, which are not as strongly charged as glass, can still support PEMs. The charge of the final layer of PEMs on weakly-charged surfaces becomes as high as that of PEMs on strongly-charged surfaces, as long as the PEM has sufficiently-many layers. For example, PEM formation on O<sub>2</sub>-plasma treated silicone rubber has been demonstrated by the present inventors. This means that all the advantages of the glass/PEM/biotin/Streptavidin/biotin-DNA surface chemistry can be applied to other substrates.

Although the above discussion describes the immobilization of polynucleotide templates by attachment to the surface of flow channels or the surface of reaction chambers disposed along flow channels, other methods of template immobilization can also be employed in the methods of the present invention. In some methods, the templates can be attached to microbeads, which can be arranged within the microfluidic system. For instance, commercially-available latex microspheres with pre-defined surface chemistry can be used. The polynucleotide templates can be attached either before or after the microbeads are inducted into the microfluidic system. Attachment of template before beads are added allows a reduction in system complexity and setup time (as many templates can be attached to different aliquots of beads simultaneously). Attachment of template to beads in situ can allow easier manipulation of surface chemistry (as bead surface chemistry can be manipulated in bulk and externally to the microfluidic device). Beads should be held in place within the flow system for this technique to be effective. Methods to achieve this include, e.g., flowing the beads into orifices too small for them to flow through (where they would become "wedged in"), the creation of "microscreens" (i.e. barriers in the channel with apertures too small for beads to pass through), and insertion of the beads into hollows in the channels where they are affixed by simple Van der Waals forces.

#### C. Primer extension reaction

Once templates are immobilized to the surfaces of synthesis channels, primer extension reactions are performed (E. D. Hyman, *Anal. Biochem.*, 174, p. 423, 1988). If part of the template sequence is known, a specific primer can be constructed and hybridized to the template. Alternatively, a linker can be ligated to the template of unknown sequence in order to allow for hybridization of a primer. The primer can be hybridized to the template before or after immobilization of the template to the surface of the synthesis channel.

In some methods, the primer is extended by a nucleic acid polymerase in the presence of a single type of labeled nucleotide. Label is incorporated into the template/primer complex only if the labeled nucleotide added to the reaction is complementary to the nucleotide on the template adjacent the 3' end of the primer. The template is subsequently washed to remove any unincorporated label, and the presence of any incorporated label is determined. A radioactive label can be determined by counting or any other method known in the art, while fluorescent labels can be induced to fluoresce, e.g., by excitation.

In some applications of the present invention, a combination of labeled and non-labeled nucleotides are used in the analysis. Because there are multiple copies of each template molecule immobilized on the surface of the synthesis channel, a small percentage of labeled nucleotides is sufficient for detection by a detection device (see below). For example,  
5 for fluorescently labeled nucleotides, the percentage of labeled nucleotide can be less than 20%, less than 10%, less than 5%, less than 1%, less than 0.1%, less than 0.01%, or even less than 0.001% of the total labeled and unlabeled nucleotides for each type of the nucleotides.

### 1. Labeled Nucleotides

10 In some methods, at least one and usually all types of the deoxyribonucleotides (dATP, dTTP, dGTP, dCTP, dUTP/dTTP) or nucleotides (ATP, UTP, GTP, and CTP) are labeled. Various labels which are easily detected include radioactive labels, optically detectable labels, spectroscopic labels and the like. Preferably, fluorescent labels are used. The different types of nucleotides can be labeled with the same kind of  
15 labels. Alternatively, a different kind of label can be used to label each different type of nucleotide.

In some methods, fluorescent labels are used. the fluorescent label can be selected from any of a number of different moieties. The preferred moiety is a fluorescent group for which detection is quite sensitive. For example, fluorescein- or rhodamine-labeled  
20 nucleotide triphosphates are available (e.g., from NEN DuPont).

Fluorescently labeled nucleotide triphosphates can also be made by various fluorescence-labeling techniques, e.g., as described in Kambara et al. (1988) "Optimization of Parameters in a DNA Sequenator Using Fluorescence Detection," Bio/Technol. 6:816-821; Smith et al. (1985) Nucl. Acids Res, 13:2399-2412; and Smith et al. (1986) Nature 321:674-  
25 679. Fluorescent labels exhibiting particularly high coefficients of destruction can also be useful in destroying nonspecific background signals.

### 2. Blocking agents:

30 In some methods during the primer extension step, a chain elongation inhibitor can be employed in the reaction (see, e.g., Dower et al., U.S. Patent No. 5,902,723. Chain elongation inhibitors are nucleotide analogues which either are chain terminators which prevent further addition by the polymerase of nucleotides to the 3' end of the chain by becoming incorporated into the chain themselves. In some methods, the chain elongation inhibitors are dideoxynucleotides. Where the chain elongation inhibitors are incorporated

into the growing polynucleotide chain, they should be removed after incorporation of the labeled nucleotide has been detected, in order to allow the sequencing reaction to proceed using different labeled nucleotides. Some 3' to 5' exonucleases, e.g., exonuclease III, are able to remove dideoxynucleotides.

5                   Other than chain elongation inhibitors, a blocking agent or blocking group can be employed on the 3' moiety of the deoxyribose group of the labeled nucleotide to prevent nonspecific incorporation. Optimally, the blocking agent should be removable under mild conditions (e.g., photosensitive, weak acid labile, or weak base labile groups), thereby allowing for further elongation of the primer strand with a next synthetic cycle. If the  
10 blocking agent also contains the fluorescent label, the dual blocking and labeling functions are achieved without the need for separate reactions for the separate moieties. For example, the labeled nucleotide can be labeled by attachment of a fluorescent dye group to the 3' moiety of the deoxyribose group, and the label is removed by cleaving the fluorescent dye from the nucleotide to generate a 3' hydroxyl group. The fluorescent dye is preferably linked  
15 to the deoxyribose by a linker arm which is easily cleaved by chemical or enzymatic means.

Examples of blocking agents include, among others, light sensitive groups such as 6-nitroveratryloxycarbonyl (NVOC), 2-nitrobenzyloxycarbonyl (NBOC), . $\alpha$ ., $\alpha$ -dimethyl-dimethoxybenzyloxycarbonyl (DDZ), 5-bromo-7-nitroindoliny, o-hydroxy-2-methyl cinnamoyl, 2-oxymethylene anthraquinone, and t-butyl oxycarbonyl (TBOC). Other  
20 blocking reagents are discussed, e.g., in U.S. Ser. No. 07/492,462; Patchornik (1970) J. Amer. Chem. Soc. 92:6333; and Amit et al. (1974) J. Org. Chem. 39:192. Nucleotides possessing various labels and blocking groups can be readily synthesized. Labeling moieties are attached at appropriate sites on the nucleotide using chemistry and conditions as described, e.g., in Gait (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press,  
25 Oxford.

### 3. Polymerases

Depending on the template, either RNA polymerase or DNA polymerases can be used in the primer extension. For analysis of DNA templates, many DNA polymerases are  
30 available. Examples of suitable DNA polymerases include, but are not limited to, Sequenase 2.0.RTM., T4 DNA polymerase or the Klenow fragment of DNA polymerase 1, or Vent polymerase. In some methods, polymerases which lack 3'  $\rightarrow$  5' exonuclease activity can be used (e.g., T7 DNA polymerase (Amersham) or Klenow fragment of DNA polymerase I

(New England Biolabs)). In some methods, when it is desired that the polymerase have proof-reading activity, polymerases lacking 3' → 5' exonuclease activity are not used. In some methods, thermostable polymerases such as ThermoSequenase™ (Amersham) or Taquenase™ (ScienTech, St Louis, MO) are used.

5           The nucleotides used in the methods should be compatible with the selected polymerase. Procedures for selecting suitable nucleotide and polymerase combinations can be adapted from Ruth et al. (1981) *Molecular Pharmacology* 20:415-422; Kutateladze, T., et al. (1984) *Nuc. Acids Res.*, 12:1671-1686; Chidgeavadze, Z., et al. (1985) *FEBS Letters*, 183:275-278.

10           The polymerase can be stored in a separate reservoir in the apparatus and flowed into the synthesis channels prior to each extension reaction cycle. The enzyme can also be stored together with the other reaction agents (e.g., the nucleotide triphosphates). Alternatively, the polymerase can be immobilized onto the surface of the synthesis channel along with the polynucleotide template.

#### 15           4. Removal of blocking group and labels

          By repeating the incorporation and label detection steps until incorporation is detected, the nucleotide on the template adjacent the 3' end of the primer can be identified. Once this has been achieved, the label should be removed before repeating the process to  
20       discover the identity of the next nucleotide. Removal of the label can be effected by removal of the labeled nucleotide using a 3'-5' exonuclease and subsequent replacement with an unlabeled nucleotide. Alternatively, the labeling group can be removed from the nucleotide. In a further alternative, where the label is a fluorescent label, it is possible to neutralize the label by bleaching it with radiation. Photobleaching can be performed according to methods,  
25       e.g., as described in Jacobson et al., "International Workshop on the Application of Fluorescence Photobleaching Techniques to Problems in Cell Biology", *Federation Proceedings*, 42:72-79, 1973; Okabe et al., *J Cell Biol* 120:1177-86, 1993; and Close et al., *Radiat Res* 53:349-57, 1973.

30           If chain terminators or 3' blocking groups have been used, these should be removed before the next cycle can take place. 3' blocking groups can be removed by chemical or enzymatic cleavage of the blocking group from the nucleotide. For example, chain terminators are removed with a 3'-5' exonuclease, e.g., exonuclease III. Once the label

and terminators/blocking groups have been removed, the cycle is repeated to discover the identity of the next nucleotide.

Removal of the blocking groups can be unnecessary if the labels are removable. In this approach, the chains incorporating the blocked nucleotides are permanently terminated and no longer participate in the elongation processes. So long as these blocked nucleotides are also removed from the labeling process, a small percentage of permanent loss in each cycle can also be tolerated.

In some methods, other than labeled nucleotides, nucleotide incorporation is monitored by detection of pyrophosphate release (see, e.g., WO98/13523, WO98/28440, and Ronaghi et al., *Science* 281:363, 1998). For example, a pyrophosphate-detection enzyme cascade is included in the reaction mixture in order to produce a chemoluminescent signal. Also, instead of deoxynucleotides or dideoxynucleotides, nucleotide analogues are used which are capable of acting as substrates for the polymerase but incapable of acting as substrates for the pyrophosphate-detection enzyme. Pyrophosphate is released upon incorporation of a deoxynucleotide or dideoxynucleotide, which can be detected enzymatically. This method employs no wash steps, instead relying on continual addition of reagents.

#### D. Detection of incorporated signals and scanning system

##### 1. Optical detection

Methods for visualizing single molecules of DNA labeled with an intercalating dye include, e.g., fluorescence microscopy as described in Houseal et al., *Biophysical Journal* 56: 507, 1989. While usually signals from a plurality of molecules are to be detected with the sequencing methods of the present invention, fluorescence from single fluorescent dye molecules can also be detected. For example, a number of methods are available for this purpose (see, e.g., Nie et al., *Science* 266: 1013, 1994; Funatsu et al., *Nature* 374: 555, 1995; Mertz et al., *Optics Letters* 20: 2532, 1995; and Unger et al., *Biotechniques* 27:1008, 1999). Even the fluorescent spectrum and lifetime of a single molecule before it photobleaches can be measured (Macklin et al., *Science* 272: 255, 1996). Standard detectors such as a photomultiplier tube or avalanche photodiode can be used. Full field imaging with a two stage image intensified CCD camera can also be used (Funatsu et al., *supra*).

The detection system for the signal or label can also depend upon the label used, which can be defined by the chemistry available. For optical signals, a combination of an optical fiber or charged couple device (CCD) can be used in the detection step. In those

circumstances where the matrix is itself transparent to the radiation used, it is possible to have an incident light beam pass through the substrate with the detector located opposite the substrate from the polynucleotides. For electromagnetic labels, various forms of spectroscopy systems can be used. Various physical orientations for the detection system are available and discussion of important design parameters is provided, e.g., in Jovin, Adv. in Biochem. Biophys.

Incorporated signals can be detected by scanning the synthesis channels. The synthesis channels can be scanned simultaneously or serially, depending on the scanning method used. The signals can be scanned using a CCD camera (TE/CCD512SF, Princeton Instruments, Trenton, N.J.) with suitable optics (Ploem, J. S., in *Fluorescent and Luminescent Probes for Biological Activity*, Mason, T. W., Ed., Academic Press, London, pp. 1-11, 1993), such as described in Yershov et al. (Proc. Natl. Acad. Sci. 93:4913, 1996), or can be imaged by TV monitoring (Khrapko et al., DNA Sequencing 1:375, 1991). For radioactive signals (e.g.,  $^{32}\text{P}$ ), a phosphorimager device can be used (Johnston et al., Johnston, R. F., et al., Electrophoresis 11:355, 1990; and Drmanac et al., Drmanac, R., et al., Electrophoresis 13:566, 1992). These methods are particularly useful to achieve simultaneous scanning of multiple probe-regions.

For fluorescence labeling, the synthesis channels can be serially scanned one by one or row by row using a fluorescence microscope apparatus, such as described in U.S. Patent Nos. 6,094,274, 5,902,723, 5,424,186, and 5,091,652. In some methods, standard low-light level cameras, such as a SIT and image intensified CCD camera, are employed (see, Funatsu et al., *Nature* 374, 555, 1995). An ICCD can be preferable to a cooled CCD camera because of its better time resolution. These devices are commercially available (e.g., from Hamamatsu).

Alternatively, only the intensifier unit from Hamamatsu or DEP are used and incorporated into other less expensive or home built cameras. If necessary, the intensifier can be cooled. For example, CCD camera can be purchased from Phillips, who offer a low priced, low noise (40 electron readout noise per pixel) model. A home built camera allows greater flexibility in the choice of components and a higher performance device. The advantage of using a camera instead of an avalanche photodiode is that one can image the whole field of view. This extra spatial information allows the development of new noise reduction techniques. For example, one can use the fact that signals are expected from certain spatial locations (i.e. where the polynucleotide template is attached) in order to reject noise.

In some applications, fluorescent excitation is exerted with a Q-switched frequency doubled Nd YAG laser, which has a KHz repetition rate, allowing many samples to be taken per second. For example, a wavelength of 532 nm is ideal for the excitation of rhodamine. It is a standard device that has been used in the single molecule detection scheme (Smith et al.,  
5 *Science* 253:1122, 1992). A pulsed laser allows time resolved experiments, which are useful for rejecting extraneous noise. In some methods, excitation can be performed with a mercury lamp and signals from the incorporated nucleotides can be detected with an inexpensive CCD camera (see, e.g., Unger et al., *Biotechniques* 27:1008, 1999).

The scanning system should be able to reproducibly scan the synthesis  
10 channels in the apparatuses. Where appropriate, e.g., for a two dimensional substrate where the synthesis channels are localized to positions thereon, the scanning system should positionally define the synthesis channels attached thereon to a reproducible coordinate system. It is important that the positional identification of synthesis channels be repeatable in successive scan steps.

Various scanning systems can be employed in the apparatuses of the present  
15 invention. For example, electrooptical scanning devices described in, e.g., U.S. Pat. No. 5,143,854, are suitable for use with the apparatuses of the present invention. The system could exhibit many of the features of photographic scanners, digitizers or even compact disk reading devices. For example, a model no. PM500-A1 x-y translation table manufactured by  
20 Newport Corporation can be attached to a detector unit. The x-y translation table is connected to and controlled by an appropriately programmed digital computer such as an IBM PC/AT or AT compatible computer. The detection system can be a model no. R943-02 photomultiplier tube manufactured by Hamamatsu, attached to a preamplifier, e.g., a model no. SR440 manufactured by Stanford Research Systems, and to a photon counter, e.g., an  
25 SR430 manufactured by Stanford Research System, or a multichannel detection device. Although a digital signal can usually be preferred, there can be circumstances where analog signals would be advantageous.

The stability and reproducibility of the positional localization in scanning  
determine, to a large extent, the resolution for separating closely positioned polynucleotide  
30 clusters on a 2 dimensional substrate. Since the successive monitoring at a given position depends upon the ability to map the results of a reaction cycle to its effect on a positionally mapped cluster of polynucleotides, high resolution scanning is preferred. As the resolution increases, the upper limit to the number of possible polynucleotides which can be sequenced on a single matrix also increases. Crude scanning systems can resolve only on the order of



1000  $\mu\text{m}$ , refined scanning systems can resolve on the order of 100  $\mu\text{m}$ , more refined systems can resolve on the order of about 10  $\mu\text{m}$ , and with optical magnification systems a resolution on the order of 1.0  $\mu\text{m}$  is available. The limitations on the resolution can be diffraction limited and advantages can arise from using shorter wavelength radiation for fluorescent scanning steps. However, with increased resolution, the time required to fully scan a matrix can increased and a compromise between speed and resolution can be selected. Parallel detection devices which provide high resolution with shorter scan times are applicable where multiple detectors are moved in parallel.

In some applications, resolution often is not so important and sensitivity is emphasized. However, the reliability of a signal can be pre-selected by counting photons and continuing to count for a longer period at positions where intensity of signal is lower. Although this decreases scan speed, it can increase reliability of the signal determination. Various signal detection and processing algorithms can be incorporated into the detection system. In some methods, the distribution of signal intensities of pixels across the region of signal are evaluated to determine whether the distribution of intensities corresponds to a time positive signal.

## 2. Non-optical detection

Other than fluorescently labeled nucleotides and optical detection devices, other methods of detecting nucleotide incorporation are also contemplated in the present invention, including the use of mass spectrometry to analyze the reaction products, the use of radiolabeled nucleotides, and detection of reaction products with "wired enzymes".

In some methods, mass spectrometry is employed to detect nucleotide incorporation in the primer extension reaction. A primer extension reaction consumes a nucleotide triphosphate, adds a single base to the primer/template duplex, and produces pyrophosphate as a by-product. Mass spectrometry can be used to detect pyrophosphate in the wash stream after a nucleotide has been incubated with the template and polymerase. The absence of pyrophosphate indicates that the nucleotide was not incorporated, whereas the presence of pyrophosphate indicates incorporation. Detection based on pyrophosphate release have been described, e.g., in WO98/13523, WO98/28440, and Ronaghi et al., Science 281:363, 1998.

In some methods, radiolabeled nucleotides are used. Nucleotides can be radiolabeled either in the sugar, the base, or the triphosphate group. To detect radioactivity, small radioactivity sensor can be incorporated in the substrate on which the microfluidic chip

is mounted. A CCD pixel, for instance, serves as a good detector for some radioactive decay processes. Radiolabeling of the sugar or base produces an additive signal: each incorporation increases the amount of radiolabel in the primer-template duplex. If the nucleotide is labeled in the portion that is released as pyrophosphate (e.g. dNTP with  $\beta$ - or  $\gamma$ - $^{32}\text{P}$ ), the radioactive pyrophosphate can be detected in the wash stream. This radioactivity level is not additive, but rather binary for each attempted nucleotide addition, so subsequent addition poses no read length limit. Due to the small reagent consumption and contained nature of microfluidics, the total radioactivity used in such a system is relatively minimal, and containment is relatively simple.

In some methods, non-optical detection of pyrophosphate release makes use of "wired redox enzymes" as described, e.g., in Heller et al., *Analytical Chemistry* 66:2451-2457, 1994; and Ohara et al., *Analytical Chemistry* 65:3512-3517, 1993. Briefly, enzymes are covalently linked to a hydrogel matrix containing redox active groups capable of transporting charge. The analyte to be detected is either acted on directly by a redox enzyme (either releasing or consuming electrons) or consumed as a reagent in an enzymatic cascade that produces a substrate that is reduced or oxidized by a redox enzyme. The production or consumption of electrons is detected at a metal electrode in contact with the hydrogel. For the detection of pyrophosphate, an enzymatic cascade using pyrophosphatase, maltose phosphorylase, and glucose oxidase can be employed. Pyrophosphatase converts pyrophosphate into phosphate; maltose phosphorylase converts maltose (in the presence of phosphate) to glucose 1-phosphate and glucose. Then, glucose oxidase converts the glucose to gluconolactone and  $\text{H}_2\text{O}_2$ ; this final reaction is the redox step which gives rise to a detectable current at the electrode. Glucose sensors based on this principle are well known in the art, and enzymatic cascades as described here have been demonstrated previously. Other enzymatic cascades besides the specific example given here are also contemplated the present invention. This type of detection scheme allows direct electrical readout of nucleotide incorporation at each reaction chamber, allowing easy parallelization.

#### E. Fluorescent photobleaching sequencing

In some methods, polynucleotide sequences are analyzed with a fluorescent photobleaching method. In this methods, fluorescently labeled nucleotides are used in the primer extension. Signals from the incorporated nucleotides are removed by photobleaching before next extension cycle starts.

The polynucleotide templates can be prepared as described above (e.g., cloning in single-stranded M13 plasmid). Biotinylated templates are attached to surface of the synthesis channel that has been pretreated with the PEM technique as discussed above. After the primed, single stranded DNA is immobilized to the synthesis channel in the flow cell. A polymerase and one nucleotide triphosphate, e.g. dATP, are flowed into the flow cell. A high fidelity polymerase with no exonuclease proofreading ability is preferred. In some methods, only a fraction (e.g., less than 10%, 5%, 1%, 0.1%, 0.01%, or 0.001%) of each type of the nucleotide triphosphates is fluorescently labeled (e.g., rhodamine-labeled nucleotide triphosphates from NEN DuPont). For example, if the first base of DNA sequence following the primer is T, then the polymerase incorporates the dATP's and some fraction of the DNA molecules become fluorescently labeled. If the first base is anything else, no fluorescent molecules become incorporated. The reagents are then flowed out of the flow cell, and the fluorescence of the DNA is measured. If no fluorescence is detected, the procedure is repeated with one of the other nucleotide triphosphates. If fluorescence is detected, the identity of the first base in the sequence has been determined. The fluorescence signal is photobleached and extinguished before the procedure is then repeated for the next base in the template sequence.

The fluorescence can be excited with, e.g., a Q-switched frequency doubled Nd YAG laser (Smith et al., *Science* 253: 1122, 1992). This is a standard device used in the single molecule detection scheme that measures the fluorescent spectrum and lifetime of a single molecule before it photobleached. It has a kHz repetition rate, allowing many samples to be taken per second. The wavelength can be, e.g., 532 nm that is ideal for the excitation of rhodamine. A pulsed laser allows time resolved experiments and is useful for rejecting extraneous noise.

Detection of the incorporated label can be performed with a standard low-light level cameras, such as a SIT or a image intensified CCD camera (Funatsu et al, supra). An Intensified CCD (ICCD) camera is preferable to a cooled CCD camera because of its better time resolution. These devices are available from, e.g., Hamamatsu. However, because these cameras are extremely expensive, a detection device can be made by building just the intensifier unit from Hamamatsu into a CCD camera. Optionally, the intensifier can be cooled. The CCD camera is available from Phillips, e.g., a low priced, low noise model (40 electron readout noise per pixel). A customarily built camera allows greater flexibility in the choice of components and a higher performance device. The advantage of using a camera instead of an avalanche photodiode is that the whole field of view can be imaged. This extra

spatial information allows the development of new noise reduction techniques. For example, the fact that signals are expected from certain spatial locations (i.e. where the DNA is attached) can be used to reject noise.

## 5 F. Other considerations

A combination of factors affect the read length and throughput of the sequencing analysis according to the present invention. First, all of the unincorporated labeled nucleotides should be removed from the synthesis channel or reaction chamber after each cycle. Since only relatively small number of incorporated dye molecules are to be  
10 detected, the reagent exchange should be leave substantially fewer unincorporated labeled nucleotides than the number of nucleotides to be detected. Second, the rate of reagent exchange is limited by fluid mechanic considerations. Turbulent flow should be avoided in order to preserve effective reagent exchange, and the fluid flow shear forces should be small enough in order to not break the DNA or dislocate the enzyme. Third, the kinetics of  
15 nucleotide incorporation and enzyme-DNA complex formation should be considered.

The present invention teaches how to determine acceptable flow rate of fluids in the apparatuses. According to the invention, flow rate in the apparatuses with microfabricated flow channels having a depth of 100  $\mu\text{m}$  is typically 0.1-1 cm/sec. For microfabricated flow channels with a depth of 10  $\mu\text{m}$ , the flow rate is usually in the range of  
20 1-10 cm/sec. Fluid flow in the apparatuses remains laminar as long as the Reynolds number  $R = \rho v \iota / \eta \ll 1$ , where  $\rho$  is the density of the fluid,  $v$  is the velocity,  $\iota$  is the dimension of the chamber, and  $\eta$  is the viscosity (see, e.g., Landau et al., Fluid Mechanics, Pergamon Press, New York, 1989). The limiting velocity is in the order of 1 cm/sec for a 100 $\mu\text{m}$  channel depth. For microchannels with a depth of 10  $\mu\text{m}$ , the limit is 10 cm/sec.

25 The ultimate limit on the rate at which fluid can be exchanged is determined by the effect of drag and shear flows on the polynucleotide template and the polymerase. The velocity profile of constrained flow is parabolic ( $v(\tau) - v_{\text{ave}}(1 - (\tau/R)^2)$ ), causing a shear force. Single molecule experiments with double stranded DNA have shown that one can apply forces of up to  $F = 50$  pN without breaking or causing irreversible damage to DNA (see, e.g.,  
30 Smith et al., Science 271: 795, 1996; and Cluzel et al., Science 271: 792, 1996), and a similar order of magnitude is expected for single stranded DNA. The drag coefficient of DNA  $\alpha = 6\pi R_5$  can be estimated from the radius of gyration  $R_5 = 0.3\mu\text{m}$ . Then the maximum fluid velocity allowed is determined by solving the equation:

$$v_{max}(R-R_g)=F/\alpha$$

The maximum average velocity before shearing of DNA becomes a problem is 140 cm/sec.

Another consideration is to prevent the polymerase from falling off the template or becoming damaged. With RNA polymerase, it has been shown that the stalling force for RNA polymerase, at which it might receive irreversible damage, is 14 pN (Yin et al., *Science* 270:1653, 1995). Since one the drag coefficient of a DNA polymerase can be estimated from its size, a similar calculation as for the DNA shear leads to a maximum velocity of 500 cm/sec.

The time to remove all of the free nucleotides can be calculated by including the effects of diffusion into hydrodynamic calculation of the fluid flow. There are a great variety of products available, including electronic switching valves with very small dead volumes. For example, a six port valve from Upchurch with electric motor from Thar Designs has a dead volume of 2  $\mu$ l and switching time of 166 msec. Combined with 0.0025" I.D. tubing and the estimated 1  $\mu$ l capacity of the microfabricated flow cell, 4  $\mu$ l of material should be exchanged for each step in the process. A syringe or peristaltic pump can give very high flow rates, the limiting factor is low Reynolds number. The inverse rate constant to get rid of all of the nucleotides is

$$\tau=(LR/v_{ave})^{2/3} (D)^{-1/3}$$

where L is the linear dimension of the device and D is the diffusion constant of the nucleotides. Plugging in approximate numbers gives a time of  $\tau=15$  sec. To reduce the nucleotide concentration from in the order of millimolar to 1 labeled nucleotide per detection region, which is a reduction of approximately  $10^{-7}$ . The amount of time to completely flush the device is  $\ln(10^{-7})\tau=4$  minutes.

For apparatuses with microfabricated flow channel depth of 10  $\mu$ m and microfabricated valves incorporated on chip, the dead volume is reduced and throughput increased. The valves can provide an essentially zero dead volume and 10 msec switching time. This and the reduced dimensions of the device leads to a drastic increase of throughput: the time to flush the reagents (e.g., nucleotides) from the system is reduced to 0.8 sec. The overall throughput is approximately 1 base per second. Table 1 summarizes the various factors affecting throughput of apparatuses with microfabricated flow channels having a depth of 100  $\mu$ m or 10  $\mu$ m.

TABLE I. Parameters affecting throughput of the sequencing apparatuses

	I	II
Channel depth ( $\mu\text{m}$ )	100	10
Dead Volume ( $\mu\text{l}$ )	4	$10^{-3}$
Turbulence vel. (cm/sec)	1	10
DNA Shear (cm/sec)	140	14
Polymerase stall (cm/sec)	1000	100
Reagent exchange (sec)	240	0.8

Note that in apparatuses I, the limiting factor is the fluid velocity that causes turbulent flow. In apparatuses II, shear forces on the DNA also becoming limiting. The reagent exchange time is expected to improve by a factor of 100 in apparatuses II.

5

The DNA polymerases can fall off of the DNA. If enzyme is replenished, it takes time for the enzyme to find and bind to a free DNA site. This could affect throughput of the apparatuses. The attrition rate of the polymerase can be determined according to methods described in the art. For example, using the kinetics of the T4 DNA polymerase as nominal values (Taylor et al., *J. Phys. D. Appl. Phys.* 24:1443, 1991), an on-rate of  $11 \mu\text{M}^{-1} \text{sec}^{-1}$  was obtained. Hence a  $1 \mu\text{M}$  concentration of enzyme gives an on rate of  $11 \text{sec}^{-1}$ , and after 1 second, 99.3% of the DNA have polymerase bound. In the absence of nucleotides (for example, during fluorescence measurement) the polymerase falls off of the DNA with a time constant of  $0.2 \text{sec}^{-1}$  (Yin et al., *Science* 270:1653, 1995). In other words, after 5 seconds without nucleotides, this can become a source of attrition. It can be compensated for by the addition of fresh polymerase with every sequencing cycle of the device.

10  
15

For the high throughput device (e.g., apparatus II in Table I), the reagent exchange is fast enough that polymerase falling off has no significant effect on the throughput. Also, the rate of incorporation of nucleotides by the polymerase is typically about 300 bases per second. This is not a rate limiting factor for the device throughput.

20

Read length of the sequencing analysis can be affected by various factors. However, photobleaching is unlikely to cause any chemical changes to the polynucleotide template that prevent the attachment of the next base. During the photobleaching, the dye molecule is held off from the DNA on a linker arm, and it gives off so few photons that the interaction cross section is negligible. Any attrition of the labeled nucleotides also does not present any significant problem. The statistics of the photobleaching scheme are robust enough to allow sequencing to continue in spite of any attrition of the labeled nucleotides. For example, if 0.1% of the bases are labeled, then after 3000 bases the attrition is 95% if incorporation of a labeled nucleotide terminates strand extension completely. In other word, if one starts with  $10^5$  molecules, then on the first base one expects to get a fluorescent signal

25  
30

from 100 dye molecules. By the 3000th base, the signal is reduced to only 5 dye molecules. This is still detectable, since the lower limit of detection is one dye molecule.

It should also be noted that the attrition are discussed above is an extreme scenario because there is little reason to expect total attrition for each incorporated base.

5 Attrition is more likely to occur when the polymerase incorporates two successive labeled nucleotides. If 1% of the bases are labeled, the chance of incorporating two labeled nucleotides next to each other is  $1\%^2=0.01\%$ . Then the attrition rate after 3000 bases is 25%. In other words, the signal only decreases by 25% by the 3000th base. Thus, attrition does not cause a problem in this sequencing scheme.

10 Another factor that can affect read length is misincorporation. If the DNA polymerase is starved for the proper nucleotide, it can incorporate the wrong nucleotide. Misincorporation efficiencies have been measured to be three to five orders of magnitude below the efficiency for proper nucleotide incorporation (Echols et al., *Ann. Rev. Biochem* 60:477, 1991). Misincorporation can be minimized by only exposing the DNA polymerase-  
15 DNA complexes to nucleotides for as much time as is needed to incorporate the proper nucleotide. For a high fidelity DNA polymerase, misincorporation happens with a frequency of about  $10^{-4}$ . If dephasing due to misincorporation is treated as total attrition, the attrition is only 25% after 3 kb, i.e, the signal is reduced to 75% of its original. Thus, misincorporation does not hinder a 3 kb or perhaps longer read length.

20 Many modifications and variations of this invention can be made without departing from its spirit and scope. The specific embodiments described herein are for illustration only and are not intended to limit the invention in any way.

25 All publications, figures, patents and patent applications cited herein are hereby expressly incorporated by reference for all purposes to the same extent as if each was so individually denoted.